ISOLATED HUMAN TRANSPORTER PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS, AND USES THEREOF

FIELD OF THE INVENTION

The present invention is in the field of transporter proteins that are related to the Na+-independent transporter subfamily, recombinant DNA molecules, and protein production. The present invention specifically provides novel peptides and proteins that effect ligand transport and nucleic acid molecules encoding such peptide and protein molecules, all of which are useful in the development of human therapeutics and diagnostic compositions and methods.

BACKGROUND OF THE INVENTION

Transporters

Transporter proteins regulate many different functions of a cell, including cell proliferation, differentiation, and signaling processes, by regulating the flow of molecules such as ions and macromolecules, into and out of cells. Transporters are found in the plasma membranes of virtually every cell in eukaryotic organisms. Transporters mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of molecules and ion across cell membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, transporters, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Transporters are generally classified by structure and the type of mode of action. In addition, transporters are sometimes classified by the molecule type that is transported, for example, sugar transporters, chlorine channels, potassium channels, etc. There may be many classes of channels for transporting a single type of molecule (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters: Receptor and transporter nomenclature supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 (1997).

The following general classification scheme is known in the art and is followed in the present discoveries.

Channel-type transporters. Transmembrane channel proteins of this class are ubiquitously found in the membranes of all types of organisms from bacteria to higher

eukaryotes. Transport systems of this type catalyze facilitated diffusion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism. These channel proteins usually consist largely of a-helical spanners, although b-strands may also be present and may even comprise the channel. However, outer membrane porin-type channel proteins are excluded from this class and are instead included in class 9.

Carrier-type transporters. Transport systems are included in this class if they utilize a carrier-mediated process to catalyze uniport (a single species is transported by facilitated diffusion), antiport (two or more species are transported in opposite directions in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy) and/or symport (two or more species are transported together in the same direction in a tightly coupled process, not coupled to a direct form of energy other than chemiosmotic energy).

Pyrophosphate bond hydrolysis-driven active transporters. Transport systems are included in this class if they hydrolyze pyrophosphate or the terminal pyrophosphate bond in ATP or another nucleoside triphosphate to drive the active uptake and/or extrusion of a solute or solutes. The transport protein may or may not be transiently phosphorylated, but the substrate is not phosphorylated.

PEP-dependent, phosphoryl transfer-driven group translocators. Transport systems of the bacterial phosphoenolpyruvate:sugar phosphotransferase system are included in this class. The product of the reaction, derived from extracellular sugar, is a cytoplasmic sugar-phosphate.

Decarboxylation-driven active transporters. Transport systems that drive solute (e.g., ion) uptake or extrusion by decarboxylation of a cytoplasmic substrate are included in this class.

Oxidoreduction-driven active transporters. Transport systems that drive transport of a solute (e.g., an ion) energized by the flow of electrons from a reduced substrate to an oxidized substrate are included in this class.

Light-driven active transporters. Transport systems that utilize light energy to drive transport of a solute (e.g., an ion) are included in this class.

Mechanically-driven active transporters. Transport systems are included in this class if they drive movement of a cell or organelle by allowing the flow of ions (or other solutes) through the membrane down their electrochemical gradients.

Outer-membrane porins (of b-structure). These proteins form transmembrane pores or channels that usually allow the energy independent passage of solutes across a membrane. The transmembrane portions of these proteins consist exclusively of b-strands that form a b-barrel. These porin-type proteins are found in the outer membranes of Gram-negative bacteria, mitochondria and eukaryotic plastids.

Methyltransferase-driven active transporters. A single characterized protein currently falls into this category, the Na+-transporting methyltetrahydromethanopterin:coenzyme M methyltransferase.

Non-ribosome-synthesized channel-forming peptides or peptide-like molecules. These molecules, usually chains of L- and D-amino acids as well as other small molecular building blocks such as lactate, form oligomeric transmembrane ion channels. Voltage may induce channel formation by promoting assembly of the transmembrane channel. These peptides are often made by bacteria and fungi as agents of biological warfare.

Non-Proteinaceous Transport Complexes. Ion conducting substances in biological membranes that do not consist of or are not derived from proteins or peptides fall into this category.

Functionally characterized transporters for which sequence data are lacking. Transporters of particular physiological significance will be included in this category even though a family assignment cannot be made.

Putative transporters in which no family member is an established transporter. Putative transport protein families are grouped under this number and will either be classified elsewhere when the transport function of a member becomes established, or will be eliminated from the TC classification system if the proposed transport function is disproven. These families include a member or members for which a transport function has been suggested, but evidence for such a function is not yet compelling.

Auxiliary transport proteins. Proteins that in some way facilitate transport across one or more biological membranes but do not themselves participate directly in transport are included in this class. These proteins always function in conjunction with one or more transport proteins. They may provide a function connected with energy coupling to transport, play a structural role in complex formation or serve a regulatory function.

Transporters of unknown classification. Transport protein families of unknown classification are grouped under this number and will be classified elsewhere when the transport process and energy coupling mechanism are characterized. These families include

at least one member for which a transport function has been established, but either the mode of transport or the energy coupling mechanism is not known.

Ion channels

An important type of transporter is the ion channel. Ion channels regulate many different cell proliferation, differentiation, and signaling processes by regulating the flow of ions into and out of cells. Ion channels are found in the plasma membranes of virtually every cell in eukaryotic organisms. Ion channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ion across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, ion channels, such as chloride channels, also regulate organelle pH. For a review, see Greger, R. (1988) Annu. Rev. Physiol. 50:111-122.

Ion channels are generally classified by structure and the type of mode of action. For example, extracellular ligand gated channels (ELGs) are comprised of five polypeptide subunits, with each subunit having 4 membrane spanning domains, and are activated by the binding of an extracellular ligand to the channel. In addition, channels are sometimes classified by the ion type that is transported, for example, chlorine channels, potassium channels, etc. There may be many classes of channels for transporting a single type of ion (a detailed review of channel types can be found at Alexander, S.P.H. and J.A. Peters (1997). Receptor and ion channel nomenclature supplement. Trends Pharmacol. Sci., Elsevier, pp. 65-68 and http://www-biology.ucsd.edu/~msaier/transport/toc.html.

There are many types of ion channels based on structure. For example, many ion channels fall within one of the following groups: extracellular ligand-gated channels (ELG), intracellular ligand-gated channels (ILG), inward rectifying channels (INR), intercellular (gap junction) channels, and voltage gated channels (VIC). There are additionally recognized other channel families based on ion-type transported, cellular location and drug sensitivity. Detailed information on each of these, their activity, ligand type, ion type, disease association, drugability, and other information pertinent to the present invention, is well known in the art.

Extracellular ligand-gated channels, ELGs, are generally comprised of five polypeptide subunits, Unwin, N. (1993), Cell 72: 31-41; Unwin, N. (1995), Nature 373: 37-43; Hucho, F., et al., (1996) J. Neurochem. 66: 1781-1792; Hucho, F., et al., (1996) Eur. J. Biochem. 239: 539-557; Alexander, S.P.H. and J.A. Peters (1997), Trends Pharmacol. Sci.,

Elsevier, pp. 4-6; 36-40; 42-44; and Xue, H. (1998) J. Mol. Evol. 47: 323-333. Each subunit has 4 membrane spanning regions: this serves as a means of identifying other members of the ELG family of proteins. ELG bind a ligand and in response modulate the flow of ions. Examples of ELG include most members of the neurotransmitter-receptor family of proteins, e.g., GABAI receptors. Other members of this family of ion channels include glycine receptors, ryandyne receptors, and ligand gated calcium channels.

The Voltage-gated Ion Channel (VIC) Superfamily

Proteins of the VIC family are ion-selective channel proteins found in a wide range of bacteria, archaea and eukaryotes Hille, B. (1992), Chapter 9: Structure of channel proteins; Chapter 20: Evolution and diversity. In: Ionic Channels of Excitable Membranes, 2nd Ed., Sinaur Assoc. Inc., Pubs., Sunderland, Massachusetts; Sigworth, F.J. (1993), Quart. Rev. Biophys. 27: 1-40; Salkoff, L. and T. Jegla (1995), Neuron 15: 489-492; Alexander, S.P.H. et al., (1997), Trends Pharmacol. Sci., Elsevier, pp. 76-84; Jan, L.Y. et al., (1997), Annu. Rev. Neurosci. 20: 91-123; Doyle, D.A, et al., (1998) Science 280: 69-77; Terlau, H. and W. Stühmer (1998), Naturwissenschaften 85: 437-444. They are often homo- or heterooligomeric structures with several dissimilar subunits (e.g., a1-a2-d-b Ca^{2+} channels, ab_1b_2 Na^+ channels or $(a)_4$-b K^+ channels), but the channel and the primary receptor is usually associated with the a (or a1) subunit. Functionally characterized members are specific for K^+, Na^+ or Ca^{2+}. The K^+ channels usually consist of homotetrameric structures with each a-subunit possessing six transmembrane spanners (TMSs). The a1 and a subunits of the Ca^{2+} and Na^+ channels, respectively, are about four times as large and possess 4 units, each with 6 TMSs separated by a hydrophilic loop, for a total of 24 TMSs. These large channel proteins form heterotetra-unit structures equivalent to the homotetrameric structures of most K^+ channels. All four units of the Ca^{2+} and Na^+ channels are homologous to the single unit in the homotetrameric K^+ channels. Ion flux via the eukaryotic channels is generally controlled by the transmembrane electrical potential (hence the designation, voltage-sensitive) although some are controlled by ligand or receptor binding.

Several putative K^+-selective channel proteins of the VIC family have been identified in prokaryotes. The structure of one of them, the KcsA K^+ channel of *Streptomyces lividans*, has been solved to 3.2 Å resolution. The protein possesses four identical subunits, each with two transmembrane helices, arranged in the shape of an inverted teepee or cone. The cone cradles the "selectivity filter" P domain in its outer end. The narrow selectivity filter is only 12 Å long, whereas the remainder of the channel is wider and lined with hydrophobic

residues. A large water-filled cavity and helix dipoles stabilize K^+ in the pore. The selectivity filter has two bound K^+ ions about 7.5 Å apart from each other. Ion conduction is proposed to result from a balance of electrostatic attractive and repulsive forces.

In eukaryotes, each VIC family channel type has several subtypes based on pharmacological and electrophysiological data. Thus, there are five types of Ca^{2+} channels (L, N, P, Q and T). There are at least ten types of K^+ channels, each responding in different ways to different stimuli: voltage-sensitive [Ka, Kv, Kvr, Kvs and Ksr], Ca^{2+}-sensitive [BK_{Ca}, IK_{Ca} and SK_{Ca}] and receptor-coupled [K_M and K_{ACh}]. There are at least six types of Na^+ channels (I, II, III, μ1, H1 and PN3). Tetrameric channels from both prokaryotic and eukaryotic organisms are known in which each a-subunit possesses 2 TMSs rather than 6, and these two TMSs are homologous to TMSs 5 and 6 of the six TMS unit found in the voltage-sensitive channel proteins. KcsA of *S. lividans* is an example of such a 2 TMS channel protein. These channels may include the K_{Na} (Na^+-activated) and K_{Vol} (cell volume-sensitive) K^+ channels, as well as distantly related channels such as the Tok1 K^+ channel of yeast, the TWIK-1 inward rectifier K^+ channel of the mouse and the TREK-1 K^+ channel of the mouse. Because of insufficient sequence similarity with proteins of the VIC family, inward rectifier K^+ IRK channels (ATP-regulated; G-protein-activated) which possess a P domain and two flanking TMSs are placed in a distinct family. However, substantial sequence similarity in the P region suggests that they are homologous. The b, g and d subunits of VIC family members, when present, frequently play regulatory roles in channel activation/deactivation.

The Epithelial Na^+ Channel (ENaC) Family

The ENaC family consists of over twenty-four sequenced proteins (Canessa, C.M., et al., (1994), Nature 367: 463-467, Le, T. and M.H. Saier, Jr. (1996), Mol. Membr. Biol. 13: 149-157; Garty, H. and L.G. Palmer (1997), Physiol. Rev. 77: 359-396; Waldmann, R., et al., (1997), Nature 386: 173-177; Darboux, I., et al., (1998), J. Biol. Chem. 273: 9424-9429; Firsov, D., et al., (1998), EMBO J. 17: 344-352; Horisberger, J.-D. (1998). Curr. Opin. Struc. Biol. 10: 443-449). All are from animals with no recognizable homologues in other eukaryotes or bacteria. The vertebrate ENaC proteins from epithelial cells cluster tightly together on the phylogenetic tree: voltage-insensitive ENaC homologues are also found in the brain. Eleven sequenced *C. elegans* proteins, including the degenerins, are distantly related to the vertebrate proteins as well as to each other. At least some of these proteins form part of a mechano-transducing complex for touch sensitivity. The homologous *Helix*

aspersa (FMRF-amide)-activated Na^+ channel is the first peptide neurotransmitter-gated ionotropic receptor to be sequenced.

Protein members of this family all exhibit the same apparent topology, each with N- and C-termini on the inside of the cell, two amphipathic transmembrane spanning segments, and a large extracellular loop. The extracellular domains contain numerous highly conserved cysteine residues. They are proposed to serve a receptor function.

Mammalian ENaC is important for the maintenance of Na^+ balance and the regulation of blood pressure. Three homologous ENaC subunits, alpha, beta, and gamma, have been shown to assemble to form the highly Na^+-selective channel. The stoichiometry of the three subunits is $alpha_2$, beta1, gamma1 in a heterotetrameric architecture.

The Glutamate-gated Ion Channel (GIC) Family of Neurotransmitter Receptors

Members of the GIC family are heteropentameric complexes in which each of the 5 subunits is of 800-1000 amino acyl residues in length (Nakanishi, N., et al, (1990), Neuron 5: 569-581; Unwin, N. (1993), Cell 72: 31-41; Alexander, S.P.H. and J.A. Peters (1997) Trends Pharmacol. Sci., Elsevier, pp. 36-40). These subunits may span the membrane three or five times as putative a-helices with the N-termini (the glutamate-binding domains) localized extracellularly and the C-termini localized cytoplasmically. They may be distantly related to the ligand-gated ion channels, and if so, they may possess substantial b-structure in their transmembrane regions. However, homology between these two families cannot be established on the basis of sequence comparisons alone. The subunits fall into six subfamilies: a, b, g, d, e and z.

The GIC channels are divided into three types: (1) a-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA)-, (2) kainate- and (3) N-methyl-D-aspartate (NMDA)-selective glutamate receptors. Subunits of the AMPA and kainate classes exhibit 35-40% identity with each other while subunits of the NMDA receptors exhibit 22-24% identity with the former subunits. They possess large N-terminal, extracellular glutamate-binding domains that are homologous to the periplasmic glutamine and glutamate receptors of ABC-type uptake permeases of Gram-negative bacteria. All known members of the GIC family are from animals. The different channel (receptor) types exhibit distinct ion selectivities and conductance properties. The NMDA-selective large conductance channels are highly permeable to monovalent cations and Ca^{2+}. The AMPA- and kainate-selective ion channels are permeable primarily to monovalent cations with only low permeability to Ca^{2+}.

The Chloride Channel (ClC) Family

The ClC family is a large family consisting of dozens of sequenced proteins derived from Gram-negative and Gram-positive bacteria, cyanobacteria, archaea, yeast, plants and animals (Steinmeyer, K., et al., (1991), Nature 354: 301-304; Uchida, S., et al., (1993), J. Biol. Chem. 268: 3821-3824; Huang, M.-E., et al., (1994), J. Mol. Biol. 242: 595-598; Kawasaki, M., et al, (1994), Neuron 12: 597-604; Fisher, W.E., et al., (1995), Genomics. 29:598-606; and Foskett, J.K. (1998), Annu. Rev. Physiol. 60: 689-717). These proteins are essentially ubiquitous, although they are not encoded within genomes of *Haemophilus influenzae*, *Mycoplasma genitalium*, and *Mycoplasma pneumoniae*. Sequenced proteins vary in size from 395 amino acyl residues *(M. jannaschii)* to 988 residues (man). Several organisms contain multiple ClC family paralogues. For example, *Synechocystis* has two paralogues, one of 451 residues in length and the other of 899 residues. *Arabidopsis thaliana* has at least four sequenced paralogues, (775-792 residues), humans also have at least five paralogues (820-988 residues), and *C. elegans* also has at least five (810-950 residues). There are nine known members in mammals, and mutations in three of the corresponding genes cause human diseases. *E. coli*, *Methanococcus jannaschii* and *Saccharomyces cerevisiae* only have one ClC family member each. With the exception of the larger *Synechocystis* paralogue, all bacterial proteins are small (395-492 residues) while all eukaryotic proteins are larger (687-988 residues). These proteins exhibit 10-12 putative transmembrane a-helical spanners (TMSs) and appear to be present in the membrane as homodimers. While one member of the family, *Torpedo* ClC-O, has been reported to have two channels, one per subunit, others are believed to have just one.

All functionally characterized members of the ClC family transport chloride, some in a voltage-regulated process. These channels serve a variety of physiological functions (cell volume regulation; membrane potential stabilization; signal transduction; transepithelial transport, etc.). Different homologues in humans exhibit differing anion selectivities, i.e., ClC4 and ClC5 share a $NO_3^- > Cl^- > Br^- > I^-$ conductance sequence, while ClC3 has an $I^- > Cl^-$ selectivity. The ClC4 and ClC5 channels and others exhibit outward rectifying currents with currents only at voltages more positive than +20mV.

Animal Inward Rectifier K^+ Channel (IRK-C) Family

IRK channels possess the "minimal channel-forming structure" with only a P domain, characteristic of the channel proteins of the VIC family, and two flanking transmembrane

spanners (Shuck, M.E., et al., (1994), J. Biol. Chem. 269: 24261-24270; Ashen, M.D., et al., (1995), Am. J. Physiol. 268: H506-H511; Salkoff, L. and T. Jegla (1995), Neuron 15: 489-492; Aguilar-Bryan, L., et al., (1998), Physiol. Rev. 78: 227-245; Ruknudin, A., et al., (1998), J. Biol. Chem. 273: 14165-14171). They may exist in the membrane as homo- or heterooligomers. They have a greater tendency to let K^+ flow into the cell than out. Voltage-dependence may be regulated by external K^+, by internal Mg^{2+}, by internal ATP and/or by G-proteins. The P domains of IRK channels exhibit limited sequence similarity to those of the VIC family, but this sequence similarity is insufficient to establish homology. Inward rectifiers play a role in setting cellular membrane potentials, and the closing of these channels upon depolarization permits the occurrence of long duration action potentials with a plateau phase. Inward rectifiers lack the intrinsic voltage sensing helices found in VIC family channels. In a few cases, those of Kir1.1a and Kir6.2, for example, direct interaction with a member of the ABC superfamily has been proposed to confer unique functional and regulatory properties to the heteromeric complex, including sensitivity to ATP. The SUR1 sulfonylurea receptor (spQ09428) is the ABC protein that regulates the Kir6.2 channel in response to ATP, and CFTR may regulate Kir1.1a. Mutations in SUR1 are the cause of familial persistent hyperinsulinemic hypoglycemia in infancy (PHHI), an autosomal recessive disorder characterized by unregulated insulin secretion in the pancreas.

ATP-gated Cation Channel (ACC) Family

Members of the ACC family (also called P2X receptors) respond to ATP, a functional neurotransmitter released by exocytosis from many types of neurons (North, R.A. (1996), Curr. Opin. Cell Biol. 8: 474-483; Soto, F., M. Garcia-Guzman and W. Stühmer (1997), J. Membr. Biol. 160: 91-100). They have been placed into seven groups ($P2X_1$ - $P2X_7$) based on their pharmacological properties. These channels, which function at neuron-neuron and neuron-smooth muscle junctions, may play roles in the control of blood pressure and pain sensation. They may also function in lymphocyte and platelet physiology. They are found only in animals.

The proteins of the ACC family are quite similar in sequence (>35% identity), but they possess 380-1000 amino acyl residues per subunit with variability in length localized primarily to the C-terminal domains. They possess two transmembrane spanners, one about 30-50 residues from their N-termini, the other near residues 320-340. The extracellular receptor domains between these two spanners (of about 270 residues) are well conserved with numerous conserved glycyl and cysteyl residues. The hydrophilic C-termini vary in

length from 25 to 240 residues. They resemble the topologically similar epithelial Na^+ channel (ENaC) proteins in possessing (a) N- and C-termini localized intracellularly, (b) two putative transmembrane spanners, (c) a large extracellular loop domain, and (d) many conserved extracellular cysteyl residues. ACC family members are, however, not demonstrably homologous with them. ACC channels are probably hetero- or homomultimers and transport small monovalent cations (Me^+). Some also transport Ca^{2+}; a few also transport small metabolites.

The Ryanodine-Inositol 1,4,5-triphosphate Receptor Ca^{2+} Channel (RIR-CaC) Family

Ryanodine (Ry)-sensitive and inositol 1,4,5-triphosphate (IP3)-sensitive Ca^{2+}-release channels function in the release of Ca^{2+} from intracellular storage sites in animal cells and thereby regulate various Ca^{2+} -dependent physiological processes (Hasan, G. et al., (1992) Development 116: 967-975; Michikawa, T., et al., (1994), J. Biol. Chem. 269: 9184-9189; Tunwell, R.E.A., (1996), Biochem. J. 318: 477-487; Lee, A.G. (1996) *Biomembranes,* Vol. 6, Transmembrane Receptors and Channels (A.G. Lee, ed.), JAI Press, Denver, CO., pp 291-326; Mikoshiba, K., et al., (1996) J. Biochem. Biomem. 6: 273-289). Ry receptors occur primarily in muscle cell sarcoplasmic reticular (SR) membranes, and IP3 receptors occur primarily in brain cell endoplasmic reticular (ER) membranes where they effect release of Ca^{2+} into the cytoplasm upon activation (opening) of the channel.

The Ry receptors are activated as a result of the activity of dihydropyridine-sensitive Ca^{2+} channels. The latter are members of the voltage-sensitive ion channel (VIC) family. Dihydropyridine-sensitive channels are present in the T-tubular systems of muscle tissues.

Ry receptors are homotetrameric complexes with each subunit exhibiting a molecular size of over 500,000 daltons (about 5,000 amino acyl residues). They possess C-terminal domains with six putative transmembrane a -helical spanners (TMSs). Putative pore-forming sequences occur between the fifth and sixth TMSs as suggested for members of the VIC family. The large N-terminal hydrophilic domains and the small C-terminal hydrophilic domains are localized to the cytoplasm. Low resolution 3-dimensional structural data are available. Mammals possess at least three isoforms that probably arose by gene duplication and divergence before divergence of the mammalian species. Homologues are present in humans and *Caenorabditis elegans.*

IP_3 receptors resemble Ry receptors in many respects. (1) They are homotetrameric complexes with each subunit exhibiting a molecular size of over 300,000 daltons (about 2,700 amino acyl residues). (2) They possess C-terminal channel domains that are

homologous to those of the Ry receptors. (3) The channel domains possess six putative TMSs and a putative channel lining region between TMSs 5 and 6. (4) Both the large N-terminal domains and the smaller C-terminal tails face the cytoplasm. (5) They possess covalently linked carbohydrate on extracytoplasmic loops of the channel domains. (6) They have three currently recognized isoforms (types 1, 2, and 3) in mammals which are subject to differential regulation and have different tissue distributions.

IP_3 receptors possess three domains: N-terminal IP_3-binding domains, central coupling or regulatory domains and C-terminal channel domains. Channels are activated by IP_3 binding, and like the Ry receptors, the activities of the IP_3 receptor channels are regulated by phosphorylation of the regulatory domains, catalyzed by various protein kinases. They predominate in the endoplasmic reticular membranes of various cell types in the brain but have also been found in the plasma membranes of some nerve cells derived from a variety of tissues.

The channel domains of the Ry and IP_3 receptors comprise a coherent family that in spite of apparent structural similarities, do not show appreciable sequence similarity of the proteins of the VIC family. The Ry receptors and the IP_3 receptors cluster separately on the RIR-CaC family tree. They both have homologues in *Drosophila.* Based on the phylogenetic tree for the family, the family probably evolved in the following sequence: (1) A gene duplication event occurred that gave rise to Ry and IP_3 receptors in invertebrates. (2) Vertebrates evolved from invertebrates. (3) The three isoforms of each receptor arose as a result of two distinct gene duplication events. (4) These isoforms were transmitted to mammals before divergence of the mammalian species.

The Organellar Chloride Channel (O-ClC) Family

Proteins of the O-ClC family are voltage-sensitive chloride channels found in intracellular membranes but not the plasma membranes of animal cells (Landry, D, et al., (1993), J. Biol. Chem. 268: 14948-14955; Valenzuela, Set al., (1997), J. Biol. Chem. 272: 12575-12582; and Duncan, R.R., et al., (1997), J. Biol. Chem. 272: 23880-23886).

They are found in human nuclear membranes, and the bovine protein targets to the microsomes, but not the plasma membrane, when expressed in *Xenopus laevis* oocytes. These proteins are thought to function in the regulation of the membrane potential and in transepithelial ion absorption and secretion in the kidney. They possess two putative transmembrane a-helical spanners (TMSs) with cytoplasmic N- and C-termini and a large luminal loop that may be glycosylated. The bovine protein is 437 amino acyl residues in

length and has the two putative TMSs at positions 223-239 and 367-385. The human nuclear protein is much smaller (241 residues). A *C. elegans* homologue is 260 residues long.

The present invention has a substantial similarity to rat small intestine Na+-independent transporter for aromatic amino acids that designated as TAT1 (T-type amino acid transporter 1).

System T was originally characterized in human erythrocytes. It transports aromatic amino acids in a Na^+-independent manner. Although it was once proposed that system T is a variant of system L which shows Na^+-independent transport of neutral amino acids including aromatic amino acids, system T is distinct in that it accepts *N*-methyl amino acids whereas system L does not. Therefore, it is reasonable to assume that transporters subserving system T would belong to a different family with distinct mechanisms of substrate recognition.

The Na+-independent transporter is Na+-independent and low-affinity transport of aromatic amino acids such as tryptophan, tyrosine, and phenylalanine (Km values: approximately 5 mm), consistent with the properties of classical amino acid transport system T. TAT1 accepted some variations of aromatic side chains because it interacted with amino acid-related compounds such as l-DOPA and 3-O-methyl-DOPA. TAT1 recognizes amino acid substrates as anions, because TAT1 accepted N-methyl- and N-acetyl-derivatives of aromatic amino acids but did not accept their methylesters. Consistent with this, TAT1 exhibited sequence similarity (approximately 30% identity at the amino acid level) to H+/monocarboxylate transporters. Different from H+/monocarboxylate transporters, however, TAT1 was not coupled with the H+ transport but it mediates an electroneutral facilitated diffusion. In rat small intestine TAT1 immunoreactivity was detected in the basolateral membrane of the epithelial cells suggesting its role in the transepithelial transport of aromatic amino acids. For a further review of Na+-independent transporter, see Kim et al., J Biol Chem 2001 May 18;276(20):17221-8.

Transporter proteins, particularly members of the Na+-independent transporter subfamily, are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown transport proteins. The present invention advances the state of the art by providing previously unidentified human transport proteins.

SUMMARY OF THE INVENTION

The present invention is based in part on the identification of amino acid sequences of human transporter peptides and proteins that are related to the Na+-independent transporter subfamily, as well as allelic variants and other mammalian orthologs thereof. These unique peptide sequences, and nucleic acid sequences that encode these peptides, can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins, and serve as targets for the development of human therapeutic agents that modulate transporter activity in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues.

DESCRIPTION OF THE FIGURE SHEETS

FIGURE 1 provides the nucleotide sequence of a cDNA molecule or transcript sequence that encodes the transporter protein of the present invention. (SEQ ID NO:1) In addition structure and functional information is provided, such as ATG start, stop and tissue distribution, where available, that allows one to readily determine specific uses of inventions based on this molecular sequence. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues.

FIGURE 2 provides the predicted amino acid sequence of the transporter of the present invention. (SEQ ID NO:2) In addition structure and functional information such as protein family, function, and modification sites is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence.

FIGURE 3 provides genomic sequences that span the gene encoding the transporter protein of the present invention. (SEQ ID NO:3) In addition structure and functional information, such as intron/exon structure, promoter location, etc., is provided where available, allowing one to readily determine specific uses of inventions based on this molecular sequence. 94 SNPs, including 10 indels, have been identified in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

DETAILED DESCRIPTION OF THE INVENTION

General Description

The present invention is based on the sequencing of the human genome. During the sequencing and assembly of the human genome, analysis of the sequence information revealed previously unidentified fragments of the human genome that encode peptides that share structural and/or sequence homology to protein/peptide/domains identified and characterized within the art as being a transporter protein or part of a transporter protein and are related to the Na+-independent transporter subfamily. Utilizing these sequences, additional genomic sequences were assembled and transcript and/or cDNA sequences were isolated and characterized. Based on this analysis, the present invention provides amino acid sequences of human transporter peptides and proteins that are related to the Na+-independent transporter subfamily, nucleic acid sequences in the form of transcript sequences, cDNA sequences and/or genomic sequences that encode these transporter peptides and proteins, nucleic acid variation (allelic information), tissue distribution of expression, and information about the closest art known protein/peptide/domain that has structural or sequence homology to the transporter of the present invention.

In addition to being previously unknown, the peptides that are provided in the present invention are selected based on their ability to be used for the development of commercially important products and services. Specifically, the present peptides are selected based on homology and/or structural relatedness to known transporter proteins of the Na+-independent transporter subfamily and the expression pattern observed. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues.. The art has clearly established the commercial importance of members of this family of proteins and proteins that have expression patterns similar to that of the present gene. Some of the more specific features of the peptides of the present invention, and the uses thereof, are described herein, particularly in the Background of the Invention and in the annotation provided in the Figures, and/or are known within the art for each of the known Na+-independent transporter family or subfamily of transporter proteins.

Specific Embodiments

Peptide Molecules

The present invention provides nucleic acid sequences that encode protein molecules that have been identified as being members of the transporter family of proteins and are related to the Na+-independent transporter subfamily (protein sequences are provided in Figure 2, transcript/cDNA sequences are provided in Figures 1 and genomic sequences are provided in Figure 3). The peptide sequences provided in Figure 2, as well as the obvious variants described herein, particularly allelic variants as identified herein and using the information in Figure 3, will be referred herein as the transporter peptides of the present invention, transporter peptides, or peptides/proteins of the present invention.

The present invention provides isolated peptide and protein molecules that consist of, consist essentially of, or comprising the amino acid sequences of the transporter peptides disclosed in the Figure 2, (encoded by the nucleic acid molecule shown in Figure 1, transcript/cDNA or Figure 3, genomic sequence), as well as all obvious variants of these peptides that are within the art to make and use. Some of these variants are described in detail below.

As used herein, a peptide is said to be "isolated" or "purified" when it is substantially free of cellular material or free of chemical precursors or other chemicals. The peptides of the present invention can be purified to homogeneity or other degrees of purity. The level of purification will be based on the intended use. The critical feature is that the preparation allows for the desired function of the peptide, even if in the presence of considerable amounts of other components (the features of an isolated nucleic acid molecule is discussed below).

In some uses, "substantially free of cellular material" includes preparations of the peptide having less than about 30% (by dry weight) other proteins (i.e., contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins. When the peptide is recombinantly produced, it can also be substantially free of culture medium, i.e., culture medium represents less than about 20% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the peptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the transporter peptide having

less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

The isolated transporter peptide can be purified from cells that naturally express it, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. For example, a nucleic acid molecule encoding the transporter peptide is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Many of these techniques are described in detail below.

Accordingly, the present invention provides proteins that consist of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). The amino acid sequence of such a protein is provided in Figure 2. A protein consists of an amino acid sequence when the amino acid sequence is the final amino acid sequence of the protein.

The present invention further provides proteins that consist essentially of the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein consists essentially of an amino acid sequence when such an amino acid sequence is present with only a few additional amino acid residues, for example from about 1 to about 100 or so additional residues, typically from 1 to about 20 additional residues in the final protein.

The present invention further provides proteins that comprise the amino acid sequences provided in Figure 2 (SEQ ID NO:2), for example, proteins encoded by the transcript/cDNA nucleic acid sequences shown in Figure 1 (SEQ ID NO:1) and the genomic sequences provided in Figure 3 (SEQ ID NO:3). A protein comprises an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the protein. In such a fashion, the protein can be only the peptide or have additional amino acid molecules, such as amino acid residues (contiguous encoded sequence) that are naturally associated with it or heterologous amino acid residues/peptide sequences. Such a protein can have a few additional amino acid

residues or can comprise several hundred or more additional amino acids. The preferred classes of proteins that are comprised of the transporter peptides of the present invention are the naturally occurring mature proteins. A brief description of how various types of these proteins can be made/isolated is provided below.

The transporter peptides of the present invention can be attached to heterologous sequences to form chimeric or fusion proteins. Such chimeric and fusion proteins comprise a transporter peptide operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the transporter peptide. "Operatively linked" indicates that the transporter peptide and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the transporter peptide.

In some uses, the fusion protein does not affect the activity of the transporter peptide *per se*. For example, the fusion protein can include, but is not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions, MYC-tagged, HI-tagged and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant transporter peptide. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A transporter peptide-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the transporter peptide.

As mentioned above, the present invention also provides and enables obvious variants of the amino acid sequence of the proteins of the present invention, such as naturally occurring mature forms of the peptide, allelic/sequence variants of the peptides, non-naturally occurring recombinantly derived variants of the peptides, and orthologs and paralogs of the peptides. Such variants can readily be generated using art-known techniques in the fields of recombinant

nucleic acid technology and protein biochemistry. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

Such variants can readily be identified/made using molecular techniques and the sequence information disclosed herein. Further, such variants can readily be distinguished from other peptides based on sequence and/or structural homology to the transporter peptides of the present invention. The degree of homology/identity present will be based primarily on whether the peptide is a functional variant or non-functional variant, the amount of divergence present in the paralog family and the evolutionary distance between the orthologs.

To determine the percent identity of two amino acid sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of a reference sequence is aligned for comparison purposes. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62

matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al., Nucleic Acids Res. 12(1)*:387 (1984)) (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Myers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against sequence databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (*J. Mol. Biol.* 215:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (*Nucleic Acids Res.* 25(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

Full-length pre-processed forms, as well as mature processed forms, of proteins that comprise one of the peptides of the present invention can readily be identified as having complete sequence identity to one of the transporter peptides of the present invention as well as being encoded by the same genetic locus as the transporter peptide provided herein. The gene encoding the novel transporter protein of the present invention is located on a genome component that has been mapped to human chromosome 6 (as indicated in Figure 3), which is supported by multiple lines of evidence, such as STS and BAC map data.

Allelic variants of a transporter peptide can readily be identified as being a human protein having a high degree (significant) of sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by the same genetic locus as the transporter peptide provided herein. Genetic locus can readily be determined based on the genomic

information provided in Figure 3, such as the genomic sequence mapped to the reference human. The gene encoding the novel transporter protein of the present invention is located on a genome component that has been mapped to human chromosome 6 (as indicated in Figure 3), which is supported by multiple lines of evidence, such as STS and BAC map data. As used herein, two proteins (or a region of the proteins) have significant homology when the amino acid sequences are typically at least about 70-80%, 80-90%, and more typically at least about 90-95% or more homologous. A significantly homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under stringent conditions as more fully described below.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 94 SNP variants were found, including 10 indels (indicated by a "-") and 1SNPs in exons. SNPs, identified at different nucleotide positions in introns and regions 5' and 3' of the ORF, may affect control/regulatory elements.

Paralogs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide, as being encoded by a gene from humans, and as having similar activity or function. Two proteins will typically be considered paralogs when the amino acid sequences are typically at least about 60% or greater, and more typically at least about 70% or greater homology through a given region or domain. Such paralogs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions as more fully described below.

Orthologs of a transporter peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the transporter peptide as well as being encoded by a gene from another organism. Preferred orthologs will be isolated from mammals, preferably primates, for the development of human therapeutic targets and agents. Such orthologs will be encoded by a nucleic acid sequence that will hybridize to a transporter peptide encoding nucleic acid molecule under moderate to stringent conditions, as more fully described below, depending on the degree of relatedness of the two organisms yielding the proteins.

Non-naturally occurring variants of the transporter peptides of the present invention can readily be generated using recombinant techniques. Such variants include, but are not limited to

deletions, additions and substitutions in the amino acid sequence of the transporter peptide. For example, one class of substitutions are conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a transporter peptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amide residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.*, *Science 247*:1306-1310 (1990).

Variant transporter peptides can be fully functional or can lack function in one or more activities, e.g. ability to bind ligand, ability to transport ligand, ability to mediate signaling, etc. Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Figure 2 provides the result of protein analysis and can be used to identify critical domains/regions. Functional variants can also contain substitution of similar amino acids that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science 244*:1081-1085 (1989)), particularly using the results provided in Figure 2. The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as transporter activity or in assays such as an *in vitro* proliferative activity. Sites that are critical for binding partner/substrate binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.*, *J. Mol. Biol. 224*:899-904 (1992); de Vos *et al.* *Science 255*:306-312 (1992)).

The present invention further provides fragments of the transporter peptides, in addition to proteins and peptides that comprise and consist of such fragments, particularly those comprising the residues identified in Figure 2. The fragments to which the invention pertains,

however, are not to be construed as encompassing fragments that may be disclosed publicly prior to the present invention.

As used herein, a fragment comprises at least 8, 10, 12, 14, 16, or more contiguous amino acid residues from a transporter peptide. Such fragments can be chosen based on the ability to retain one or more of the biological activities of the transporter peptide or could be chosen for the ability to perform a function, e.g. bind a substrate or act as an immunogen. Particularly important fragments are biologically active fragments, peptides that are, for example, about 8 or more amino acids in length. Such fragments will typically comprise a domain or motif of the transporter peptide, e.g., active site, a transmembrane domain or a substrate-binding domain. Further, possible fragments include, but are not limited to, domain or motif containing fragments, soluble peptide fragments, and fragments containing immunogenic structures. Predicted domains and functional sites are readily identifiable by computer programs well known and readily available to those of skill in the art (e.g., PROSITE analysis). The results of one such analysis are provided in Figure 2.

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in transporter peptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art (some of these features are identified in Figure 2).

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues,

hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (*Meth. Enzymol. 182*: 626-646 (1990)) and Rattan *et al.* (*Ann. N.Y. Acad. Sci. 663*:48-62 (1992)).

Accordingly, the transporter peptides of the present invention also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature transporter peptide is fused with another compound, such as a compound to increase the half-life of the transporter peptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature transporter peptide, such as a leader or secretory sequence or a sequence for purification of the mature transporter peptide or a pro-protein sequence.

Protein/Peptide Uses

The proteins of the present invention can be used in substantial and specific assays related to the functional information provided in the Figures; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its binding partner or ligand) in biological fluids; and as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state). Where the protein binds or potentially binds to another protein or ligand (such as, for example, in a transporter-effector protein interaction or transporter-ligand interaction), the protein can be used to identify the binding partner/ligand so as to develop a system to identify inhibitors of the binding interaction. Any or all of these uses are capable of being developed into reagent grade or kit format for commercialization as commercial products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

The potential uses of the peptides of the present invention are based primarily on the source of the protein as well as the class/action of the protein. For example, transporters isolated from humans and their human/mammalian orthologs serve as targets for identifying agents for use in mammalian therapeutic applications, e.g. a human drug, particularly in modulating a biological or pathological response in a cell or tissue that expresses the transporter. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans. A large percentage of pharmaceutical agents are being developed that modulate the activity of transporter proteins, particularly members of the Na+-independent transporter subfamily (see Background of the Invention). The structural and functional information provided in the Background and Figures provide specific and substantial uses for the molecules of the present invention, particularly in combination with the expression information provided in Figure 1. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. Such uses can readily be determined using the information provided herein, that known in the art and routine experimentation.

The proteins of the present invention (including variants and fragments that may have been disclosed prior to the present invention) are useful for biological assays related to transporters that are related to members of the Na+-independent transporter subfamily. Such assays involve any of the known transporter functions or activities or properties useful for diagnosis and treatment of transporter-related conditions that are specific for the subfamily of transporters that the one of the present invention belongs to, particularly in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans. The proteins of the present invention are also useful in drug screening assays, in cell-based or cell-free systems ((Hodgson, Bio/technology, 1992, Sept 10(9);973-80). Cell-based systems can be native, i.e., cells that normally express the transporter, as a biopsy or expanded in cell culture.

Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. In an alternate embodiment, cell-based assays involve recombinant host cells expressing the transporter protein.

The polypeptides can be used to identify compounds that modulate transporter activity of the protein in its natural state or an altered form that causes a specific disease or pathology associated with the transporter. Both the transporters of the present invention and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the transporter. These compounds can be further screened against a functional transporter to determine the effect of the compound on the transporter activity. Further, these compounds can be tested in animal or invertebrate systems to determine activity/effectiveness. Compounds can be identified that activate (agonist) or inactivate (antagonist) the transporter to a desired degree.

Further, the proteins of the present invention can be used to screen a compound for the ability to stimulate or inhibit interaction between the transporter protein and a molecule that normally interacts with the transporter protein, e.g. a substrate or a component of the signal pathway that the transporter protein normally interacts (for example, another transporter). Such assays typically include the steps of combining the transporter protein with a candidate compound under conditions that allow the transporter protein, or fragment, to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the transporter protein and the target, such as any of the associated effects of signal transduction such as changes in membrane potential, protein phosphorylation, cAMP turnover, and adenylate cyclase activation, etc.

Candidate compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.*, *Nature 354*:82-84 (1991); Houghten *et al.*, *Nature 354*:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.*, *Cell 72*:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, $F(ab')_2$, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble fragment of the receptor that competes for ligand binding. Other candidate compounds include mutant transporters or appropriate fragments containing mutations that affect transporter function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention further includes other end point assays to identify compounds that modulate (stimulate or inhibit) transporter activity. The assays typically involve an assay of events in the signal transduction pathway that indicate transporter activity. Thus, the transport of a ligand, change in cell membrane potential, activation of a protein, a change in the expression of genes that are up- or down-regulated in response to the transporter protein dependent signal cascade can be assayed.

Any of the biological or biochemical functions mediated by the transporter can be used as an endpoint assay. These include all of the biochemical or biochemical/biological events described herein, in the references cited herein, incorporated by reference for these endpoint assay targets, and other functions known to those of ordinary skill in the art or that can be readily identified using the information provided in the Figures, particularly Figure 2. Specifically, a biological function of a cell or tissues that expresses the transporter can be assayed. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans.

Binding and/or activating compounds can also be screened by using chimeric transporter proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a ligand-binding region can be used that interacts with a different ligand then that which is recognized by the native transporter. Accordingly, a different set of signal transduction components is available as an end-point assay for activation. This allows for assays to be performed in other than the specific host cell from which the transporter is derived.

The proteins of the present invention are also useful in competition binding assays in methods designed to discover compounds that interact with the transporter (e.g. binding

partners and/or ligands). Thus, a compound is exposed to a transporter polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble transporter polypeptide is also added to the mixture. If the test compound interacts with the soluble transporter polypeptide, it decreases the amount of complex formed or activity from the transporter target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the transporter. Thus, the soluble polypeptide that competes with the target transporter region is designed to contain peptide sequences corresponding to the region of interest.

To perform cell free drug screening assays, it is sometimes desirable to immobilize either the transporter protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ^{35}S-labeled) and the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of transporter-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with binding of the protein to its target molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a transporter-binding protein and a candidate compound are incubated in the transporter protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the transporter protein target molecule, or which are reactive with transporter protein and compete with the target molecule,

as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Agents that modulate one of the transporters of the present invention can be identified using one or more of the above assays, alone or in combination. It is generally preferable to use a cell-based or cell free system first and then confirm activity in an animal or other model system. Such model systems are well known in the art and can readily be employed in this context.

Modulators of transporter protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the transporter pathway, by treating cells or tissues that express the transporter. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. These methods of treatment include the steps of administering a modulator of transporter activity in a pharmaceutical composition to a subject in need of such treatment, the modulator being identified as described herein.

In yet another aspect of the invention, the transporter proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with the transporter and are involved in transporter activity. Such transporter-binding proteins are also likely to be involved in the propagation of signals by the transporter proteins or transporter targets as, for example, downstream elements of a transporter-mediated signaling pathway. Alternatively, such transporter-binding proteins are likely to be transporter inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a transporter protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a transporter-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This

proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the transporter protein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a transporter-modulating agent, an antisense transporter nucleic acid molecule, a transporter-specific antibody, or a transporter-binding partner) can be used in an animal or other model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal or other model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The transporter proteins of the present invention are also useful to provide a target for diagnosing a disease or predisposition to disease mediated by the peptide. Accordingly, the invention provides methods for detecting the presence, or levels of, the protein (or encoding mRNA) in a cell, tissue, or organism. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. The method involves contacting a biological sample with a compound capable of interacting with the transporter protein such that the interaction can be detected. Such an assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

One agent for detecting a protein in a sample is an antibody capable of selectively binding to protein. A biological sample includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The peptides of the present invention also provide targets for diagnosing active protein activity, disease, or predisposition to disease, in a patient having a variant peptide, particularly activities and conditions that are known for other members of the family of proteins to which the present one belongs. Thus, the peptide can be isolated from a biological sample and assayed for the presence of a genetic mutation that results in aberrant peptide. This includes amino acid substitution, deletion, insertion, rearrangement, (as the result of aberrant splicing events), and

inappropriate post-translational modification. Analytic methods include altered electrophoretic mobility, altered tryptic peptide digest, altered transporter activity in cell-based or cell-free assay, alteration in ligand or antibody-binding pattern, altered isoelectric point, direct amino acid sequencing, and any other of the known assay techniques useful for detecting mutations in a protein. Such an assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

In vitro techniques for detection of peptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence using a detection reagent, such as an antibody or protein binding agent. Alternatively, the peptide can be detected *in vivo* in a subject by introducing into the subject a labeled anti-peptide antibody or other types of detection agent. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. Particularly useful are methods that detect the allelic variant of a peptide expressed in a subject and methods which detect fragments of a peptide in a sample.

The peptides are also useful in pharmacogenomic analysis. Pharmacogenomics deal with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Eichelbaum, M. (*Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 (1996)), and Linder, M.W. (*Clin. Chem.* 43(2):254-266 (1997)). The clinical outcomes of these variations result in severe toxicity of therapeutic drugs in certain individuals or therapeutic failure of drugs in certain individuals as a result of individual variation in metabolism. Thus, the genotype of the individual can determine the way a therapeutic compound acts on the body or the way the body metabolizes the compound. Further, the activity of drug metabolizing enzymes effects both the intensity and duration of drug action. Thus, the pharmacogenomics of the individual permit the selection of effective compounds and effective dosages of such compounds for prophylactic or therapeutic treatment based on the individual's genotype. The discovery of genetic polymorphisms in some drug metabolizing enzymes has explained why some patients do not obtain the expected drug effects, show an exaggerated drug effect, or experience serious toxicity from standard drug dosages. Polymorphisms can be expressed in the phenotype of the extensive metabolizer and the phenotype of the poor metabolizer. Accordingly, genetic polymorphism may lead to allelic protein variants of the transporter protein in which one or more of the transporter functions in one population is different from those in another population. The peptides thus allow a target to ascertain a genetic predisposition that can affect treatment modality. Thus, in a ligand-based

treatment, polymorphism may give rise to amino terminal extracellular domains and/or other ligand-binding regions that are more or less active in ligand binding, and transporter activation. Accordingly, ligand dosage would necessarily be modified to maximize the therapeutic effect within a given population containing a polymorphism. As an alternative to genotyping, specific polymorphic peptides could be identified.

The peptides are also useful for treating a disorder characterized by an absence of, inappropriate, or unwanted expression of the protein. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. Accordingly, methods for treatment include the use of the transporter protein or fragments.

Antibodies

The invention also provides antibodies that selectively bind to one of the peptides of the present invention, a protein comprising such a peptide, as well as variants and fragments thereof. As used herein, an antibody selectively binds a target peptide when it binds the target peptide and does not significantly bind to unrelated proteins. An antibody is still considered to selectively bind a peptide even if it also binds to other proteins that are not substantially homologous with the target peptide so long as such proteins share homology with a fragment or domain of the peptide target of the antibody. In this case, it would be understood that antibody binding to the peptide is still selective despite some degree of cross-reactivity.

As used herein, an antibody is defined in terms consistent with that recognized within the art: they are multi-subunit proteins produced by a mammalian organism in response to an antigen challenge. The antibodies of the present invention include polyclonal antibodies and monoclonal antibodies, as well as fragments of such antibodies, including, but not limited to, Fab or $F(ab')_2$, and Fv fragments.

Many methods are known for generating and/or identifying antibodies to a given target peptide. Several such methods are described by Harlow, Antibodies, Cold Spring Harbor Press, (1989).

In general, to generate antibodies, an isolated peptide is used as an immunogen and is administered to a mammalian organism, such as a rat, rabbit or mouse. The full-length protein, an antigenic peptide fragment or a fusion protein can be used. Particularly important fragments are those covering functional domains, such as the domains identified in Figure 2, and domain

of sequence homology or divergence amongst the family, such as those that can readily be identified using protein alignment methods and as presented in the Figures.

Antibodies are preferably prepared from regions or discrete fragments of the transporter proteins. Antibodies can be prepared from any region of the peptide as described herein. However, preferred regions will include those involved in function/activity and/or transporter/binding partner interaction. Figure 2 can be used to identify particularly important regions while sequence alignment can be used to identify conserved and unique sequence fragments.

An antigenic fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, at least 10, 12, 14, 16 or more amino acid residues. Such fragments can be selected on a physical property, such as fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions or can be selected based on sequence uniqueness (see Figure 2).

Detection on an antibody of the present invention can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Antibody Uses

The antibodies can be used to isolate one of the proteins of the present invention by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural protein from cells and recombinantly produced protein expressed in host cells. In addition, such antibodies are useful to detect the presence of one of the proteins of the present invention in cells or tissues to determine the pattern of expression of the protein among various tissues in an organism and over the course of normal development. Experimental data as provided in Figure 1 indicates that transporter proteins of

the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans. Further, such antibodies can be used to detect protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression. Also, such antibodies can be used to assess abnormal tissue distribution or abnormal expression during development or progression of a biological condition. Antibody detection of circulating fragments of the full length protein can be used to identify turnover.

Further, the antibodies can be used to assess expression in disease states such as in active stages of the disease or in an individual with a predisposition toward disease related to the protein's function. When a disorder is caused by an inappropriate tissue distribution, developmental expression, level of expression of the protein, or expressed/processed form, the antibody can be prepared against the normal protein. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. If a disorder is characterized by a specific mutation in the protein, antibodies specific for this mutant protein can be used to assay for the presence of the specific mutant protein.

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting expression level or the presence of aberrant sequence and aberrant tissue distribution or developmental expression, antibodies directed against the protein or relevant fragments can be used to monitor therapeutic efficacy.

Additionally, antibodies are useful in pharmacogenomic analysis. Thus, antibodies prepared against polymorphic proteins can be used to identify individuals that require modified treatment modalities. The antibodies are also useful as diagnostic tools as an immunological marker for aberrant protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for tissue typing. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as

well as in different tissues. Thus, where a specific protein has been correlated with expression in a specific tissue, antibodies that are specific for this protein can be used to identify a tissue type.

The antibodies are also useful for inhibiting protein function, for example, blocking the binding of the transporter peptide to a binding partner such as a ligand or protein binding partner. These uses can also be applied in a therapeutic context in which treatment involves inhibiting the protein's function. An antibody can be used, for example, to block binding, thus modulating (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against specific fragments containing sites required for function or against intact protein that is associated with a cell or cell membrane. See Figure 2 for structural information relating to the proteins of the present invention.

The invention also encompasses kits for using antibodies to detect the presence of a protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting protein in a biological sample; means for determining the amount of protein in the sample; means for comparing the amount of protein in the sample with a standard; and instructions for use. Such a kit can be supplied to detect a single protein or epitope or can be configured to detect one of a multitude of epitopes, such as in an antibody detection array. Arrays are described in detail below for nucleic acid arrays and similar methods have been developed for antibody arrays.

Nucleic Acid Molecules

The present invention further provides isolated nucleic acid molecules that encode a transporter peptide or protein of the present invention (cDNA, transcript and genomic sequence). Such nucleic acid molecules will consist of, consist essentially of, or comprise a nucleotide sequence that encodes one of the transporter peptides of the present invention, an allelic variant thereof, or an ortholog or paralog thereof.

As used herein, an "isolated" nucleic acid molecule is one that is separated from other nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly contiguous peptide encoding sequences and peptide encoding sequences within the same gene but separated by introns in the

genomic sequence. The important point is that the nucleic acid is isolated from remote and unimportant flanking sequences such that it can be subjected to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a transcript/cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Accordingly, the present invention provides nucleic acid molecules that consist of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists of a nucleotide sequence when the nucleotide sequence is the complete nucleotide sequence of the nucleic acid molecule.

The present invention further provides nucleic acid molecules that consist essentially of the nucleotide sequence shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule consists essentially of a nucleotide sequence when such a nucleotide sequence is present with only a few additional nucleic acid residues in the final nucleic acid molecule.

The present invention further provides nucleic acid molecules that comprise the nucleotide sequences shown in Figure 1 or 3 (SEQ ID NO:1, transcript sequence and SEQ ID NO:3, genomic sequence), or any nucleic acid molecule that encodes the protein provided in Figure 2, SEQ ID NO:2. A nucleic acid molecule comprises a nucleotide sequence when the nucleotide sequence is at least part of the final nucleotide sequence of the nucleic acid molecule. In such a fashion, the nucleic acid molecule can be only the nucleotide sequence or have additional nucleic acid residues, such as nucleic acid residues that are naturally associated with

it or heterologous nucleotide sequences. Such a nucleic acid molecule can have a few additional nucleotides or can comprise several hundred or more additional nucleotides. A brief description of how various types of these nucleic acid molecules can be readily made/isolated is provided below.

In Figures 1 and 3, both coding and non-coding sequences are provided. Because of the source of the present invention, humans genomic sequence (Figure 3) and cDNA/transcript sequences (Figure 1), the nucleic acid molecules in the Figures will contain genomic intronic sequences, 5' and 3' non-coding sequences, gene regulatory regions and non-coding intergenic sequences. In general such sequence features are either noted in Figures 1 and 3 or can readily be identified using computational tools known in the art. As discussed below, some of the non-coding regions, particularly gene regulatory elements such as promoters, are useful for a variety of purposes, e.g. control of heterologous gene expression, target for identifying gene activity modulating compounds, and are particularly claimed as fragments of the genomic sequence provided herein.

The isolated nucleic acid molecules can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature peptide (when the mature form has more than one peptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

As mentioned above, the isolated nucleic acid molecules include, but are not limited to, the sequence encoding the transporter peptide alone, the sequence encoding the mature peptide and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature peptide, with or without the additional coding sequences, plus additional non-coding sequences, for example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA processing (including splicing and polyadenylation signals), ribosome binding and stability of mRNA. In addition, the nucleic acid molecule may be fused to a marker sequence encoding, for example, a peptide that facilitates purification.

Isolated nucleic acid molecules can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be

double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense strand).

The invention further provides nucleic acid molecules that encode fragments of the peptides of the present invention as well as nucleic acid molecules that encode obvious variants of the transporter proteins of the present invention that are described above. Such nucleic acid molecules may be naturally occurring, such as allelic variants (same locus), paralogs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

The present invention further provides non-coding fragments of the nucleic acid molecules provided in Figures 1 and 3. Preferred non-coding fragments include, but are not limited to, promoter sequences, enhancer sequences, gene modulating sequences and gene termination sequences. Such fragments are useful in controlling heterologous gene expression and in developing screens to identify gene-modulating agents. A promoter can readily be identified as being 5' to the ATG start site in the genomic sequence provided in Figure 3.

A fragment comprises a contiguous nucleotide sequence greater than 12 or more nucleotides. Further, a fragment could at least 30, 40, 50, 100, 250 or 500 nucleotides in length. The length of the fragment will be based on its intended use. For example, the fragment can encode epitope bearing regions of the peptide, or can be useful as DNA probes and primers. Such fragments can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of gene.

A probe/primer typically comprises substantially a purified oligonucleotide or oligonucleotide pair. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 20, 25, 40, 50 or more consecutive nucleotides.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. As described in the Peptide Section, these variants comprise a nucleotide sequence

encoding a peptide that is typically 60-70%, 70-80%, 80-90%, and more typically at least about 90-95% or more homologous to the nucleotide sequence shown in the Figure sheets or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under moderate to stringent conditions, to the nucleotide sequence shown in the Figure sheets or a fragment of the sequence. Allelic variants can readily be determined by genetic locus of the encoding gene. The gene encoding the novel transporter protein of the present invention is located on a genome component that has been mapped to human chromosome 6 (as indicated in Figure 3), which is supported by multiple lines of evidence, such as STS and BAC map data.

Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 94 SNP variants were found, including 10 indels (indicated by a "-") and 1SNPs in exons. SNPs, identified at different nucleotide positions in introns and regions 5' and 3' of the ORF, may affect control/regulatory elements.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a peptide at least 60-70% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 60%, at least about 70%, or at least about 80% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65C. Examples of moderate to low stringency hybridization conditions are well known in the art.

Nucleic Acid Molecule Uses

The nucleic acid molecules of the present invention are useful for probes, primers, chemical intermediates, and in biological assays. The nucleic acid molecules are useful as a hybridization probe for messenger RNA, transcript/cDNA and genomic DNA to isolate full-length cDNA and genomic clones encoding the peptide described in Figure 2 and to isolate cDNA and genomic clones that correspond to variants (alleles, orthologs, etc.) producing the same or related peptides shown in Figure 2. 94 SNPs, including 10 indels, have been identified

in the gene encoding the transporter protein provided by the present invention and are given in Figure 3.

The probe can correspond to any sequence along the entire length of the nucleic acid molecules provided in the Figures. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. However, as discussed, fragments are not to be construed as encompassing fragments disclosed prior to the present invention.

The nucleic acid molecules are also useful as primers for PCR to amplify any given region of a nucleic acid molecule and are useful to synthesize antisense molecules of desired length and sequence.

The nucleic acid molecules are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the peptide sequences. Vectors also include insertion vectors, used to integrate into another nucleic acid molecule sequence, such as into the cellular genome, to alter *in situ* expression of a gene and/or gene product. For example, an endogenous coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The nucleic acid molecules are also useful for expressing antigenic portions of the proteins.

The nucleic acid molecules are also useful as probes for determining the chromosomal positions of the nucleic acid molecules by means of *in situ* hybridization methods. The gene encoding the novel transporter protein of the present invention is located on a genome component that has been mapped to human chromosome 6 (as indicated in Figure 3), which is supported by multiple lines of evidence, such as STS and BAC map data.

The nucleic acid molecules are also useful in making vectors containing the gene regulatory regions of the nucleic acid molecules of the present invention.

The nucleic acid molecules are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from the nucleic acid molecules described herein.

The nucleic acid molecules are also useful for making vectors that express part, or all, of the peptides.

The nucleic acid molecules are also useful for constructing host cells expressing a part, or all, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful for constructing transgenic animals expressing all, or a part, of the nucleic acid molecules and peptides.

The nucleic acid molecules are also useful as hybridization probes for determining the presence, level, form and distribution of nucleic acid expression. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans.

Accordingly, the probes can be used to detect the presence of, or to determine levels of, a specific nucleic acid molecule in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the peptides described herein can be used to assess expression and/or gene copy number in a given cell, tissue, or organism. These uses are relevant for diagnosis of disorders involving an increase or decrease in transporter protein expression relative to normal results.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA include Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a transporter protein, such as by measuring a level of a transporter-encoding nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a transporter gene has been mutated. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans.

Nucleic acid expression assays are useful for drug screening to identify compounds that modulate transporter nucleic acid expression.

The invention thus provides a method for identifying a compound that can be used to treat a disorder associated with nucleic acid expression of the transporter gene, particularly biological and pathological processes that are mediated by the transporter in cells and tissues that express it. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues. The method typically includes assaying the ability of the compound to modulate the expression of the transporter nucleic acid and thus identifying a compound that can be used to treat a disorder

characterized by undesired transporter nucleic acid expression. The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the transporter nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

The assay for transporter nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway. Further, the expression of genes that are up- or down-regulated in response to the transporter protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of transporter gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of transporter mRNA in the presence of the candidate compound is compared to the level of expression of transporter mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

The invention further provides methods of treatment, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate transporter nucleic acid expression in cells and tissues that express the transporter. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans. Modulation includes both up-regulation (i.e. activation or agonization) or down-regulation (suppression or antagonization) or nucleic acid expression.

Alternatively, a modulator for transporter nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule inhibits the transporter nucleic acid expression in the cells and tissues that

express the protein. Experimental data as provided in Figure 1 indicates expression in humans in the organs such as lung, brain and prostate etc, as well as in different tissues.

The nucleic acid molecules are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the transporter gene in clinical trials or in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing effectiveness of treatment with the compound, particularly with compounds to which a patient can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the affected cells to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the patient has not become resistant. Similarly, if the level of nucleic acid expression falls below a desirable level, administration of the compound could be commensurately decreased.

The nucleic acid molecules are also useful in diagnostic assays for qualitative changes in transporter nucleic acid expression, and particularly in qualitative changes that lead to pathology. The nucleic acid molecules can be used to detect mutations in transporter genes and gene expression products such as mRNA. The nucleic acid molecules can be used as hybridization probes to detect naturally occurring genetic mutations in the transporter gene and thereby to determine whether a subject with the mutation is at risk for a disorder caused by the mutation. Mutations include deletion, addition, or substitution of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion or transposition, modification of genomic DNA, such as aberrant methylation patterns or changes in gene copy number, such as amplification. Detection of a mutated form of the transporter gene associated with a dysfunction provides a diagnostic tool for an active disease or susceptibility to disease when the disease results from overexpression, underexpression, or altered expression of a transporter protein.

Individuals carrying mutations in the transporter gene can be detected at the nucleic acid level by a variety of techniques. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 94 SNP variants were found, including 10 indels (indicated by a "-") and 1 SNPs in exons. SNPs, identified at different nucleotide positions in introns and regions 5' and 3' of the ORF, may affect control/regulatory elements. The gene encoding the novel transporter protein of the present invention is located on a genome component that has been mapped to human chromosome 6 (as indicated in Figure 3), which is supported by multiple lines of evidence,

such as STS and BAC map data. Genomic DNA can be analyzed directly or can be amplified by using PCR prior to analysis. RNA or cDNA can be used in the same way. In some uses, detection of the mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.*, *Science 241*:1077-1080 (1988); and Nakazawa *et al.*, *PNAS 91*:360-364 (1994)), the latter of which can be particularly useful for detecting point mutations in the gene (see Abravaya *et al.*, *Nucleic Acids Res. 23*:675-682 (1995)). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. Deletions and insertions can be detected by a change in size of the amplified product compared to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to normal RNA or antisense DNA sequences.

Alternatively, mutations in a transporter gene can be directly identified, for example, by alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

Further, sequence-specific ribozymes (U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site. Perfectly matched sequences can be distinguished from mismatched sequences by nuclease cleavage digestion assays or by differences in melting temperature.

Sequence changes at specific locations can also be assessed by nuclease protection assays such as RNase and S1 protection or the chemical cleavage method. Furthermore, sequence differences between a mutant transporter gene and a wild-type gene can be determined by direct DNA sequencing. A variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve, C.W., (1995) *Biotechniques 19*:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.*, *Adv. Chromatogr. 36*:127-162 (1996); and Griffin *et al.*, *Appl. Biochem. Biotechnol. 38*:147-159 (1993)).

Other methods for detecting mutations in the gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes (Myers *et al.*, *Science 230*:1242 (1985)); Cotton *et al.*, *PNAS 85*:4397 (1988); Saleeba *et al.*,

Meth. Enzymol. 217:286-295 (1992)), electrophoretic mobility of mutant and wild type nucleic acid is compared (Orita *et al.*, *PNAS 86*:2766 (1989); Cotton *et al.*, *Mutat. Res. 285*:125-144 (1993); and Hayashi *et al.*, *Genet. Anal. Tech. Appl. 9*:73-79 (1992)), and movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (Myers *et al.*, *Nature 313*:495 (1985)). Examples of other techniques for detecting point mutations include selective oligonucleotide hybridization, selective amplification, and selective primer extension.

The nucleic acid molecules are also useful for testing an individual for a genotype that while not necessarily causing the disease, nevertheless affects the treatment modality. Thus, the nucleic acid molecules can be used to study the relationship between an individual's genotype and the individual's response to a compound used for treatment (pharmacogenomic relationship). Accordingly, the nucleic acid molecules described herein can be used to assess the mutation content of the transporter gene in an individual in order to select an appropriate compound or dosage regimen for treatment. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 94 SNP variants were found, including 10 indels (indicated by a "-") and 1SNPs in exons. SNPs, identified at different nucleotide positions in introns and regions 5' and 3' of the ORF, may affect control/regulatory elements.

Thus nucleic acid molecules displaying genetic variations that affect treatment provide a diagnostic target that can be used to tailor treatment in an individual. Accordingly, the production of recombinant cells and animals containing these polymorphisms allow effective clinical design of treatment compounds and dosage regimens.

The nucleic acid molecules are thus useful as antisense constructs to control transporter gene expression in cells, tissues, and organisms. A DNA antisense nucleic acid molecule is designed to be complementary to a region of the gene involved in transcription, preventing transcription and hence production of transporter protein. An antisense RNA or DNA nucleic acid molecule would hybridize to the mRNA and thus block translation of mRNA into transporter protein.

Alternatively, a class of antisense molecules can be used to inactivate mRNA in order to decrease expression of transporter nucleic acid. Accordingly, these molecules can treat a disorder characterized by abnormal or undesired transporter nucleic acid expression. This technique involves cleavage by means of ribozymes containing nucleotide sequences complementary to one or more regions in the mRNA that attenuate the ability of the mRNA to

be translated. Possible regions include coding regions and particularly coding regions corresponding to the catalytic and other functional activities of the transporter protein, such as ligand binding.

The nucleic acid molecules also provide vectors for gene therapy in patients containing cells that are aberrant in transporter gene expression. Thus, recombinant cells, which include the patient's cells that have been engineered *ex vivo* and returned to the patient, are introduced into an individual where the cells produce the desired transporter protein to treat the individual.

The invention also encompasses kits for detecting the presence of a transporter nucleic acid in a biological sample. Experimental data as provided in Figure 1 indicates that transporter proteins of the present invention are expressed in the human lung, brain, prostate, ovary, placenta, thymus, colon, and pancreas. Specifically, the protein also expressed in the tissues such as small cell carcinoma, liver, neuroblastoma cells, pooled germ cell tumors, adenocarcinoma, fibrotheoma, pooled germ cell tumors and Islets of Langerhans. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting transporter nucleic acid in a biological sample; means for determining the amount of transporter nucleic acid in the sample; and means for comparing the amount of transporter nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect transporter protein mRNA or DNA.

Nucleic Acid Arrays

The present invention further provides nucleic acid detection kits, such as arrays or microarrays of nucleic acid molecules that are based on the sequence information provided in Figures 1 and 3 (SEQ ID NOS:1 and 3).

As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, chip, glass slide, or any other suitable solid support. In one embodiment, the microarray is prepared and used according to the methods described in US Patent 5,837,832, Chee *et al.*, PCT application W095/11995 (Chee *et al.*), Lockhart, D. J. *et al.* (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. *et al.* (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference. In other embodiments, such arrays are produced by the methods described by Brown *et al.*, US Patent No. 5,807,522.

The microarray or detection kit is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray or detection kit, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in length. The microarray or detection kit may contain oligonucleotides that cover the known 5', or 3', sequence, sequential oligonucleotides that cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence. Polynucleotides used in the microarray or detection kit may be oligonucleotides that are specific to a gene or genes of interest.

In order to produce oligonucleotides to a known sequence for a microarray or detection kit, the gene(s) of interest (or an ORF identified from the contigs of the present invention) is typically examined using a computer algorithm which starts at the 5' or at the 3' end of the nucleotide sequence. Typical algorithms will then identify oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray or detection kit. The "pairs" will be identical, except for one nucleotide that preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from two to one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/251116 (Baldeschweiler *et al.*) which is incorporated herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536, 6144 or

more oligonucleotides, or any other number between two and one million which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using a microarray or detection kit, the RNA or DNA from a biological sample is made into hybridization probes. The mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray or detection kit so that the probe sequences hybridize to complementary oligonucleotides of the microarray or detection kit. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray or detection kit. The biological samples may be obtained from any bodily fluids (such as blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct sequences simultaneously. This data may be used for large-scale correlation studies on the sequences, expression patterns, mutations, variants, or polymorphisms among samples.

Using such arrays, the present invention provides methods to identify the expression of the transporter proteins/peptides of the present invention. In detail, such methods comprise incubating a test sample with one or more nucleic acid molecules and assaying for binding of the nucleic acid molecule with components within the test sample. Such assays will typically involve arrays comprising many genes, at least one of which is a gene of the present invention and or alleles of the transporter gene of the present invention. Figure 3 provides information on SNPs that have been identified in a gene encoding the transporter protein of the present invention. 94 SNP variants were found, including 10 indels (indicated by a "-") and 1SNPs in exons. SNPs, identified at different nucleotide positions in introns and regions 5' and 3' of the ORF, may affect control/regulatory elements.

Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid molecule used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or array assay formats can readily be adapted to employ the novel fragments of

the Human genome disclosed herein. Examples of such assays can be found in Chard, T, *An Introduction to Radioimmunoassay and Related Techniques,* Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. *et al., Techniques in Immunocytochemistry,* Academic Press, Orlando, FL Vol. 1 (1 982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology,* Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The test samples of the present invention include cells, protein or membrane extracts of cells. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing nucleic acid extracts or of cells are well known in the art and can be readily be adapted in order to obtain a sample that is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the nucleic acid molecules that can bind to a fragment of the Human genome disclosed herein; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound nucleic acid.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the nucleic acid probe, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound probe. One skilled in the art will readily recognize that the previously unidentified transporter gene of the present invention can be routinely identified using the sequence information disclosed herein can be readily incorporated into one of the established kit formats which are well known in the art, particularly expression arrays.

Vectors/host cells

The invention also provides vectors containing the nucleic acid molecules described herein. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, which can transport the nucleic acid molecules. When the vector is a nucleic acid molecule, the nucleic acid molecules are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the nucleic acid molecules. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the nucleic acid molecules when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the nucleic acid molecules. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the nucleic acid molecules such that transcription of the nucleic acid molecules is allowed in a host cell. The nucleic acid molecules can be introduced into the host cell with a separate nucleic acid molecule capable of affecting transcription. Thus, the second nucleic acid molecule may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the nucleic acid molecules from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself. It is understood, however, that in some embodiments, transcription and/or translation of the nucleic acid molecules can occur in a cell-free system.

The regulatory sequence to which the nucleic acid molecules described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ, the lac, TRP, and TAC promoters from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers.

Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a nucleic acid molecule. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses. Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (i.e. tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The nucleic acid molecules can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate nucleic acid molecule can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*.

Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the peptide as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the peptides. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired peptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterotransporter. Typical fusion expression vectors include pGEX (Smith *et al.*, *Gene 67*:31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, *Gene 69*:301-315 (1988)) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology 185*:60-89 (1990)).

Recombinant protein expression can be maximized in host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the nucleic acid molecule of interest can be altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.*, *Nucleic Acids Res. 20*:2111-2118 (1992)).

The nucleic acid molecules can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, *EMBO J. 6*:229-234 (1987)), pMFa (Kurjan *et al.*, *Cell 30*:933-943(1982)), pJRY88 (Schultz *et al.*, *Gene 54*:113-123 (1987)), and pYES2 (Invitrogen Corporation, San Diego, CA).

The nucleic acid molecules can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.*, *Mol. Cell Biol. 3*:2156-2165 (1983)) and the pVL series (Lucklow *et al.*, *Virology 170*:31-39 (1989)).

In certain embodiments of the invention, the nucleic acid molecules described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian

expression vectors include pCDM8 (Seed, B. *Nature 329*:840(1987)) and pMT2PC (Kaufman *et al., EMBO J. 6*:187-195 (1987)).

The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the nucleic acid molecules. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression of the nucleic acid molecules described herein. These are found for example in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the nucleic acid molecule sequences described herein, including both coding and non-coding regions. Expression of this antisense RNA is subject to each of the parameters described above in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast, other eukaryotic cells such as insect cells, and higher eukaryotic cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide sequences can be introduced on different vectors of the same cell. Similarly, the nucleic acid molecules can be introduced either alone or with other nucleic acid molecules that are not related to the nucleic acid molecules such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the nucleic acid molecule vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In the case in which viral replication is defective, replication will occur in host cells providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the nucleic acid molecules described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell- free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the peptide is desired, which is difficult to achieve with multi-transmembrane domain containing proteins such as transporters, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the peptides or heterologous to these peptides.

Where the peptide is not secreted into the medium, which is typically the case with transporters, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The peptide can then be recovered and purified by well-known purification methods including ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the peptides described herein, the peptides can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the peptides may include an initial modified methionine in some cases as a result of a host-mediated process.

Uses of vectors and host cells

The recombinant host cells expressing the peptides described herein have a variety of uses. First, the cells are useful for producing a transporter protein or peptide that can be further purified to produce desired amounts of transporter protein or fragments. Thus, host cells containing expression vectors are useful for peptide production.

Host cells are also useful for conducting cell-based assays involving the transporter protein or transporter protein fragments, such as those described above as well as other formats known in the art. Thus, a recombinant host cell expressing a native transporter protein is useful for assaying compounds that stimulate or inhibit transporter protein function.

Host cells are also useful for identifying transporter protein mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology, host cells containing the mutations are useful to assay compounds that have a desired effect on the mutant transporter protein (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native transporter protein.

Genetically engineered host cells can be further used to produce non-human transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic animal. These animals are useful for studying the function of a transporter protein and identifying and evaluating modulators of transporter protein activity. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Any of the transporter protein nucleotide sequences can be introduced as a transgene into the genome of a non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the transporter protein to particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems that allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al. PNAS 89*:6232-6236 (1992). Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al. Science 251*:1351-1355 (1991). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al. Nature 385*:810-813 (1997) and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_o phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring born of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

Transgenic animals containing recombinant cells that express the peptides described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the

various physiological factors that are present *in vivo* and that could effect ligand binding, transporter protein activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* transporter protein function, including ligand interaction, the effect of specific mutant transporter proteins on transporter protein function and ligand interaction, and the effect of chimeric transporter proteins. It is also possible to assess the effect of null mutations, that is mutations that substantially or completely eliminate one or more transporter protein functions.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

Claims

That which is claimed is:

1. An isolated peptide consisting of an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence shown in SEQ ID NO:2;

(b) an amino acid sequence of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3; and

(d) a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids.

2. An isolated peptide comprising an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence shown in SEQ ID NO:2;

(b) an amino acid sequence of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3; and

(d) a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids.

3. An isolated antibody that selectively binds to a peptide of claim 2.

4. An isolated nucleic acid molecule consisting of a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NO:2;

(b) a nucleotide sequence that encodes of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

5. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NO:2;

(b) a nucleotide sequence that encodes of an allelic variant of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NO:2, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:1 or 3;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NO:2, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

6. A gene chip comprising a nucleic acid molecule of claim 5.

7. A transgenic non-human animal comprising a nucleic acid molecule of claim 5.

8. A nucleic acid vector comprising a nucleic acid molecule of claim 5.

9. A host cell containing the vector of claim 8.

10. A method for producing any of the peptides of claim 1 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.

11. A method for producing any of the peptides of claim 2 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.

12. A method for detecting the presence of any of the peptides of claim 2 in a sample, said method comprising contacting said sample with a detection agent that specifically allows detection of the presence of the peptide in the sample and then detecting the presence of the peptide.

13. A method for detecting the presence of a nucleic acid molecule of claim 5 in a sample, said method comprising contacting the sample with an oligonucleotide that hybridizes to said nucleic acid molecule under stringent conditions and determining whether the oligonucleotide binds to said nucleic acid molecule in the sample.

14. A method for identifying a modulator of a peptide of claim 2, said method comprising contacting said peptide with an agent and determining if said agent has modulated the function or activity of said peptide.

15. The method of claim 14, wherein said agent is administered to a host cell comprising an expression vector that expresses said peptide.

16. A method for identifying an agent that binds to any of the peptides of claim 2, said method comprising contacting the peptide with an agent and assaying the contacted mixture to determine whether a complex is formed with the agent bound to the peptide.

17. A pharmaceutical composition comprising an agent identified by the method of claim 16 and a pharmaceutically acceptable carrier therefor.

18. A method for treating a disease or condition mediated by a human transporter protein, said method comprising administering to a patient a pharmaceutically effective amount of an agent identified by the method of claim 16.

19. A method for identifying a modulator of the expression of a peptide of claim 2, said method comprising contacting a cell expressing said peptide with an agent, and determining if said agent has modulated the expression of said peptide.

20. An isolated human transporter peptide having an amino acid sequence that shares at least 70% homology with an amino acid sequence shown in SEQ ID NO:2.

21. A peptide according to claim 20 that shares at least 90 percent homology with an amino acid sequence shown in SEQ ID NO:2.

22. An isolated nucleic acid molecule encoding a human transporter peptide, said nucleic acid molecule sharing at least 80 percent homology with a nucleic acid molecule shown in SEQ ID NOS:1 or 3.

23. A nucleic acid molecule according to claim 22 that shares at least 90 percent homology with a nucleic acid molecule shown in SEQ ID NOS:1 or 3.

```
   1 ATGGTGCTCT CCCAGGAGGA GCCGGACTCC GCGCGGGGCA CGAGCGAGGC
  51 GCAGCCGCTC GGCCCCGCGC CCACGGGGGC CGCTCCGCCG CCCGGCCCGG
 101 GACCCTCGGA CAGCCCCGAG GCGGCTGTCG AGAAGGTGGA GGTGGAGCTG
 151 GCGGGGCCGG CGACCGCGGA GCCCCATGAG CCCCCCGAAC CCCCCGAGGG
 201 CGGCTGGGGC TGGCTGGTGA TGCTGGCGGC CATGTGGTGC AACGGGTCGG
 251 TGTTCGGCAT CCAGAACGCT TGCGGGGTGC TCTTCGTGTC CATGCTGGAA
 301 ACCTTCGGCT CCAAAGACGA TGACAAGATG GTCTTTAAGA CAGCAGCATG
 351 GGTAGGTTCT CTCTCCATGG GGATGATTTT CTTTTGCTGC CCAATAGTCA
 401 GCGTCTTCAC AGACCTATTT GGTTGTCGGA AAACAGCTGT CGTGGGTGCT
 451 GCTGTTGGAT TTGTTGGGCT CATGTCCAGT TCTTTTGTAA GTTCCATCGA
 501 GCCTCTGTAC CTTACCTATG GAATCATATT TGCCTGCGGC TGCTCCTTTG
 551 CATACCAGCC TTCATTGGTC ATTTTGGGAC ACTATTTCAA GAAGCGCCTT
 601 GGACTGGTGA ATGGCATTGT CACTGCTGGC AGCAGTGTCT TCACAATCCT
 651 GCTGCCTTTG CTCTTAAGGG TTCTGATTGA CAGCGTGGGC CTCTTTTACA
 701 CATTGAGGGT GCTCTGCATC TTCATGTTTG TTCTCTTTCT GGCTGGCTTT
 751 ACTTACCGAC CTCTTGCTAC CAGTACCAAA GATAAAGAGA GTGGAGGTAG
 801 CGGATCCTCC CTCTTTTCCA GGAAAAAGTT CAGTCCTCCA AAAAAAATTT
 851 TCAATTTTGC CATCTTCAAG GTGACAGCTT ATGCAGTGTG GGCAGTTGGA
 901 ATACCACTTG CACTTTTTGG ATACTTTGTG CCTTATGTTC ACTTGATGAA
 951 ACATGTAAAT GAAAGATTTC AAGATGAAAA AAATAAAGAG GTTGTTCTCA
1001 TGTGCATTGG CGTCACTTCA GGAGTTGGAC GACTGCTCTT TGGCCGGATT
1051 GCAGATTATG TGCCTGGTGT GAAGAAGGTT TATCTACAGG TACTCTCCTT
1101 TTTCTTCATT GGTCTGATGT CCATGATGAT TCCTCTGTGT AGCATCTTTG
1151 GGGCCCTCAT TGCTGTGTGC CTCATCATGG GTCTCTTCGA TGGATGCTTC
1201 ATTTCCATTA TGGCTCCCAT AGCCTTTGAG TTAGTTGGTG CCCAGGATGT
1251 CTCCCAAGCA ATTGGATTTC TGCTCGGATT CATGTCTATA CCCATGACTG
1301 TTGGCCCACC CATTGCAGGG TTACTTCGTG ACAAACTGGG CTCCTATGAT
1351 GTGGCATTCT ACCTCGCTGG AGTCCCTCCC CTTATTGGAG GTGCTGTGCT
1401 TTGTTTTATC CCGTGGATCC ATAGTAAGAA GCAAAGAGAG ATCAGTAAAA
1451 CCACTGGAAA AGAAAAGATG GAGAAAATGT TGGAAAACCA GAACTCTCTG
1501 CTGTCAAGTT CATCTGGAAT GTTCAAGAAA GAATCTGACT CTATTATTTA
1551 A  (SEQ ID NO: 1)
```

FEATURES:
Start Codon: 1
Stop Codon: 1549

FIGURE 1A

HOMOLOGOUS PROTEINS:
Top BLAST Hits:
Top 10 BLAST Hits:

Sequences producing significant alignments:	Score (bits)	E Value
CRA|62000057354769 /altid=gi|14090278 /def=dbj|BAB55595.1| (AB0...	874	0.0
CRA|18000004921871 /altid=gi|5730045 /def=ref|NP_006508.1| (NM_...	527	e-148
CRA|18000004921870 /altid=gi|7513431 /def=pir||I38495 X-linked ...	527	e-148
CRA|18000005134802 /altid=gi|6677997 /def=ref|NP_033223.1| (NM_...	516	e-144
CRA|163000000492387 /altid=gi|8923981 /def=ref|NP_061063.1| (NM...	402	e-110
CRA|224000009228679 /altid=gi|17389922 /def=gb|AAH17968.1|AAH17...	399	e-109
CRA|89000000201355 /altid=gi|7299667 /def=gb|AAF54851.1| (AE003...	353	1e-95
CRA|224000007378350 /altid=gi|16768034 /def=gb|AAL28236.1| (AY0...	353	1e-95
CRA|18000005086356 /altid=gi|7449989 /def=pir||JC5507 monocarbo...	189	2e-46
CRA|18000005075554 /altid=gi|6226943 /def=sp|Q90632|MOT3_CHICK ...	188	4e-46

EST:

CRA Number	gi Number	Score	Expect
CRA|76000044050113	gi|14568965	1265 bits (638)	0.0
CRA|1000600799987	gi|6359768	1170 bits (590)	0.0
CRA|160000129712648	gi|13709362	1063 bits (536)	0.0
CRA|58000099006260	gi|12788094	1049 bits (529)	0.0
CRA|32000087643803	gi|10813242	1017 bits (513)	0.0
CRA|157000141043600	gi|13460082	973 bits (491)	0.0
CRA|160000129843301	gi|13720697	902 bits (455)	0.0
CRA|105000016327758	gi|11084182	680 bits (343)	0.0
CRA|330000005235303	gi|6989741	551 bits (278)	1e-154
CRA|3000001439560	gi|1166011	547 bits (276)	1e-153
CRA|1000684940123	gi|6569405	486 bits (245)	1e-134
CRA|1000610791502	gi|5933739	486 bits (245)	1e-134
CRA|3000001441088	gi|1164256	462 bits (233)	1e-127
CRA|160000129872523	gi|13723263	410 bits (207)	1e-112
CRA|162000043366421	gi|10877364	234 bits (118)	8e-59
CRA|3000000874607	gi|2994619	218 bits (110)	5e-54
CRA|225000013398008	gi|18086759	208 bits (105)	4e-51

EXPRESSION INFORMATION FOR MODULATORY USE:
library source:

gi Number	Organ	Tissue Type
gi|14568965	lung	small cell carcinoma
gi|6359768	(none)	liver
gi|13709362	(none)	(none)
gi|12788094	brain	neuroblastoma cells
gi|10813242	(none)	pooled germ cell tumors
gi|13460082	prostate	adenocarcinoma
gi|13720697	(none)	(none)
gi|11084182	ovary	fibrotheoma
gi|6989741	(none)	(none)
gi|1166011	placenta	(none)
gi|6569405	(none)	pooled germ cell tumors
gi|5933739	thymus, pooled	(none)
gi|1164256	placenta	(none)
gi|13723263	(none)	(none)
gi|10877364	colon	(none)
gi|2994619	pooled	(none)
gi|18086759	Pancreas	Islets of Langerhans

FIGURE 1B

```
   1 MVLSQEEPDS ARGTSEAQPL GPAPTGAAPP PGPGPSDSPE AAVEKVEVEL
  51 AGPATAEPHE PPEPPEGGWG WLVMLAAMWC NGSVFGIQNA CGVLFVSMLE
 101 TFGSKDDDKM VFKTAAWWGS LSMGMIFFCC PIVSVFTDLF GCRKTAVVGA
 151 AVGFVGLMSS SFVSSIEPLY LTYGIIFACG CSFAYQPSLV ILGHYFKKRL
 201 GLVNGIVTAG SSVFTILLPL LLRVLIDSVG LFYTLRVLCI PMFVLFLAGF
 251 TYRPLATSTK DKESGGSGSS LFSRKKFSPP KKIFNFAIFK VTAYAVWAVG
 301 IPLALFGYFV PYVHLMKHVN ERFQDEKNKE VVLMCIGVTS GVGRLLFGRI
 351 ADYVPGVKKV YLQVLSFFFI GLMSMMIPLC SIFGALIAVC LIMGLFDGCF
 401 ISIMAPIAFE LVGAQDVSQA IGFLLGPMSI PMTVGPPIAG LLRDKLGSYD
 451 VAFYLAGVPP LIGGAVLCFI PWIHSKKQRE ISKTTGKEKM EKMLENQNSL
 501 LSSSSGMFKK ESDSII (SEQ ID NO: 2)

FEATURES:
Functional domains and key regions:
PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site
              81-84   NGSV

PDOC00002 PS00002 GLYCOSAMINOGLYCAN
Glycosaminoglycan attachment site
              340-343 SGVG

PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site
Number of matches: 2
       1      275-278 KKFS
       2      509-512 KKES

PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site
Number of matches: 7
       1      10-12   SAR
       2      234-236 TLR
       3      251-253 TYR
       4      258-260 STK
       5      273-275 SRK
       6      475-477 SKK
       7      485-487 TGK

PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site
Number of matches: 6
       1      4-7     SQEE
       2      97-100  SMLE
       3      104-107 SKDD
       4      164-167 SSIE
       5      258-261 STKD
       6      485-488 TGKE

PDOC00008 PS00008 MYRISTYL
N-myristoylation site
Number of matches: 15
       1      13-18   GTSEAQ
       2      82-87   GSVFGI
       3      141-146 GCRKTA
       4      149-154 GAAVGF
       5      156-161 GLMSSS
       6      174-179 GIIFAC
       7      180-185 GCSFAY
       8      201-206 GLVNGI
       9      230-235 GLFYTL
       10     265-270 GGSGSS
       11     300-305 GIPLAL
       12     337-342 GVTSGV
       13     384-389 GALIAV
       14     394-399 GLFDGC
       15     464-469 GAVLCF
```

FIGURE 2A

```
PDOC00013 PS00013 PROKAR_LIPOPROTEIN
Prokaryotic membrane lipoprotein lipid attachment site
              169-179 LYLTYGIIFAC

PDOC00029 PS00029 LEUCINE_ZIPPER
Leucine zipper pattern
              365-386 LSFFFIGLMSMMIPLCSIFGAL

PDOC00240 PS00267 TACHYKININ
Tachykinin family signature
Number of matches: 2
       1      154-158 FVGLM
       2      369-373 FIGLM
```

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainity
1	67	87	1.683	Certain
2	117	137	2.220	Certain
3	146	166	1.923	Certain
4	169	189	1.494	Certain
5	202	222	1.761	Certain
6	232	252	1.893	Certain
7	291	311	1.874	Certain
8	330	350	0.801	Putative
9	364	384	2.458	Certain
10	387	407	1.707	Certain
11	419	439	1.966	Certain
12	453	473	1.826	Certain

```
BLAST Alignment to Top Hit:
>CRA|62000057354769 /altid=gi|14090278 /def=dbj|BAB55595.1| (AB047324)
            TAT1 [Rattus norvegicus] /org=Rattus norvegicus
            /taxon=10116 /div=ROD /dataset=nraa /length=514
          Length = 514

 Score =  874 bits (2233), Expect = 0.0
 Identities = 435/517 (84%), Positives = 463/517 (89%), Gaps = 1/517 (0%)
 Frame = +1

Query: 232  MVLSQEEPDSA-RGTSEAQPLGPAPTGAAPPPGPGPSDSPEAAVEKVEVELAGPATAEPH 408
            MV S EEP +A R T+EAQP GPAP+  AP P PGPSD  + +VEKVEVEL    +
Sbjct: 1    MVPSLEEPAAAERETNEAQPPGPAPSDDAPLPVPGPSDVSDGSVEKVEVELT--RSTGNQ 58

Query: 409  EPPEPPEGGWGWLVMLAAMWCNGSVFGIQNACGVLFVSMLETFGSKDDDKMVFKTAAWVG 588
            EPPEPPEGGWGWLVMLAAMWCNGSVFGIQNA GVLFVSMLETFG+KDDD M FK AAWVG
Sbjct: 59   EPPEPPEGGWGWLVMLAAMWCNGSVFGIQNAYGVLFVSMLETFGAKDDDNMAFK-AAWVG 117

Query: 589  SLSMGMIFFCCPIVSVFTDLFGCRKTAVVGAAVGFVGLMSSSFVSSIEPLYLTYGIIFAC 768
            SLSMGMIFFCCPIVSVFTD+FGCR+TAV+GAAVGFVGLMSSSFVSSIEPLY TYG++FAC
Sbjct: 118  SLSMGMIFFCCPIVSVFTDMFGCRRTAVLGAAVGFVGLMSSSFVSSIEPLYFTYGVVFAC 177

Query: 769  GCSFAYQPSLVILGHYFKKRLGLVNGIVTAGSSVFTILLPLLLRVLIDSVGLFYTLRVLC 948
            GCSFAYQPSLVILGHYFKKRLGLVNGIVTAGSSVFTILLPLLL  L  +VGL YTLR+LC
Sbjct: 178  GCSFAYQPSLVILGHYFKKRLGLVNGIVTAGSSVFTILLPLLLGNLTSTVGLCYTLRILC 237

Query: 949  IFMFVLFLAGFTYRPLATSTKDKESGGSGSSLFSRKKFSPPKKIFNFAIFKVTAYAVWAV 1128
            IFMFVLFLAGFTYRPL  S+K+KES  S SS FSR+K SPPKKIFNFA+FK TAYAVWA
Sbjct: 238  IFMFVLFLAGFTYRPLVPSSKEKESEDSRSSFFSRRKLSPPKKIFNFALFKETAYAVWAA 297

Query: 1129 GIPLALFGYFVPYVHLMKHVNERFQDEKNKEVVLMCIGVTSGVGRLLFGRIADYVPGVKK 1308
            GIPLALFGYFVPYVHLM HV ERF+D  NKEV+ MCIGVTSGVGRLLFGRIADY+PGVKK
Sbjct: 298  GIPLALFGYFVPYVHLMNHVKERFKDVNNKEVLFMCIGVTSGVGRLLFGRIADYLPGVKK 357

Query: 1309 VYLQVLSFFFIGLMSMMIPLCSIFGALIAVCLIMGLFDGCFISIMAPIAFELVGAQDVSQ 1488
            VYLQVLSFFFIGL SMMIPLCS+FGALIA+CLIMGLFDGCFISIMAPIAFELVG QD SQ
```

FIGURE 2B

```
Sbjct: 358  VYLQVLSFFFIGLTSMMIPLCSVFGALIALCLIMGLFDGCFISIMAPIAFELVGPQDASQ 417

Query: 1489 AIGFLLGFMSIPMTVGPPIAGLLRDKLGSYDVAFYLAGVPPLIGGAVLCFIPWIHSKKQR 1668
          AIGFLLGFMSIPMTVGPP+AGLL DKLGSYD+AFYLAG+PP IGGAVLC IPWIHSKKQR
Sbjct: 418  AIGFLLGFMSIPMTVGPPVAGLLHDKLGSYDLAFYLAGIPPFIGGAVLCLIPWIHSKKQR 477

Query: 1669 EISKTTGKEKMEKMLENQNSLLSSSSGMFKKESDSII 1779
          EISK TG EKMEKML NQ+SLLSSSSG+FKKESDSII
Sbjct: 478  EISKNTGGEKMEKMLANQSSLLSSSSGIFKKESDSII 514 (SEQ ID NO : 4)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00664	ABC transporter transmembrane region.	4.2	7.1	1
PF01027	Uncharacterized protein family	3.4	7.6	1
PF00083	Sugar (and other) transporter	2.0	7.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF00083	1/1	11	50	..	184	223	..	2.0	7.4
PF01027	1/1	121	177	..	1	59	[.	3.4	7.6
PF00664	1/1	165	232	..	1	76	[.	4.2	7.1

FIGURE 2C

```
   1 AATGGGTATT TTGTAGACTG TCCTTGATTG GGATTTGTCT AATGTTTTTC
  51 TGATGGTTAG CCAGGGATTA TGGGTTTTGG CAGGAAGACC ACAGAGGTAA
 101 AGTACCATTT TCATCACATC ATATCGGGGA TACATTATCA TCTAGTTGAG
 151 GTACTGTGTG CCATTTTTTG CACCCTAAAG TTATTTCTTC CCCCCACTCC
 201 CCCTTTCCAT CCTATACTCT TTGGAAGAAA GTTACTACGC ATACCCACAC
 251 TTAAAGAGTA AACCATTGTA CTTCACCTCC ATGAGGGAGG GAGTATGTTC
 301 ATAAAGTATT TACATTTCCT GCAGGAGAGA TTTGTCTATT CTCTCCTCAT
 351 TATTTATTTA ATCATTTACT TACATCAGTA CTGACTCGTG GATAATTCTT
 401 ACATATGTGT TTGTTTGTGT GCATGCAAAT ATATAATCGA TGTGCTTTCT
 451 TTGCCCAATA ATATGTTGTG GACAACTTTC AAAGTCAATA AATACAGATG
 501 ACCTTCAGAA CTTTTAGAGG TTTTAAAGTA AGTATCTAAT CAGTCTTCTA
 551 CCAATGTACA TTATACTTCC AAATTTTCCT TATTTCCAAC AATACTGGGG
 601 TATCATCTTC ATACATACAT TTTTGTGCAC TTATGTGCCT ATTCCTTTGT
 651 TTACTATTTT ACCCTCATTT CTAAGGCAGA TTACACTTGA GCTATGTTCC
 701 CCATTCCACA ACCAAGCACG GCTTGCTTTC CTTAGTTTAT GCTATTTTCT
 751 CTACCTGGAA TGCTCTTTTC CTGTCTTGAC CCACTGAAGT CGTATGTATG
 801 AATCAGGGCT TCGGCCAAAG GCTGTTTGCT GTAGAGGCTC TTCTACAGTG
 851 TTTGGAGAGA ATTTAAGGGA CTATCTCATC TCTCTCTTTT GTACATGTAT
 901 TATAACACAT CATCTGAGCC TCCTAGTCTC TCCCAGGACT CTTTTTTCCT
 951 ACCAGTTTAT CAACTGATAA GAGGCAGAAA CGAGATCAAT CGCACTCATC
1001 TGTGTACTCT ATCACAGTGG TGGGCACATC AAGTAACTAG CATATTTTGA
1051 CTTTGATTGA AGTGAAGAAT ACGAATAACA GAAATTAAGA AGCATCCTCA
1101 ATATTGCATA GCAGGTTACT CTTCTTTTCT TTTACATAGG ATGGCACTCC
1151 ATGCTTCAGG GAGACAGAGG AGTTGAATAC AGGTTTTAGT TTTTGTTTAA
1201 AGTGAAAACG ACTCTGATGT AGTTGAAAAG TAATGCTTTC TAGCTGTCTG
1251 TTAAAAATGT TTGTTGGTTG AAGACTTCGG AATTGCAGTC CAGTGAGGAC
1301 TGAAAATAAG CATCTTTGGT GTGCCAAATA TTCATAAGGA AATTGTATAC
1351 GAATGCAAGA GAATGGAACT GAAGTAATAA AATAAGGGCT CTGATCCTTC
1401 AGATGACTTA TTTAAGAAGC CAGGTGGCAT AACGAATCTT ACATATTATA
1451 ATTAGTACTG AGAGGTGAAT GCCAAAACAT AAAACAAACA CAATCGAGAC
1501 AATGTTAGTG TGACTGTGAC GCTGTGTCGG TGAGTTGAGG CTAACGATCC
1551 AGTGTGGCTC TCCTGAAGGC CCACCGCGCC CGCACCTAGG AGACGCGCCC
1601 CTTCTGCTCA TGCTTTGAGG CGGGGTGACC CACACATCTG TGCCCCTCTC
1651 TGAGCAGGAG GAGGCCCCGT CGCAGACGCG CGCGCAGACA GCGTCTGCCG
1701 CGGGCACCTG GGGCCGCGCG CCGCGGGGCG CCCCGCCTCC GCTCTCCGAG
1751 GCCCAATCAT CTGGAGGCTG TGGGGGCACG TCCCGCTCCC GGCCACGCCC
1801 CCAGCCGGCG GGGCGGGGGG TGCTTTTAAG AACCGGCGGC TGGCAGTGGG
1851 CTCAGTCGGG GGTGCGGGGC TGTGACCTAG AGGCTTCAGT GTCGATCCCC
1901 GAGGTGTTCG CGCGCGCCAG CTGTCCTCGC GGCCGCCTGC GCGCTGGCCG
1951 CCTGCGCGCT GCCAGCCCGC CCGCCCGCCA GGGGCTCCGC CGCCCTCGCC
2001 TCGGCCTCGT TAGCCCGCCA GGAGCCCCGC AGCTCCTCCG GGAGCCCGCT
2051 GGTAACTCGC GTCCCTCGCG CTTCTCCGGC GCCTGAGGGG CCCGCCTCGG
2101 GCCATGGTGC TCTCCCAGGA GGAGCCGGAC TCCGCGCGGG GCACGAGCGA
2151 GGCGCAGCCG CTCGGCCCCG CGCCCACGGG GGCCGCTCCG CCGCCCGGCC
2201 CGGGACCCTC GGACAGCCCC GAGGCGGCTG TCGAGAAGGT GGAGGTGGAG
2251 CTGGCGGGGC CGGCGACCGC GGAGCCCCAT GAGCCCCCCG AACCCCCCGA
2301 GGGCGGCTGG GGCTGGCTGG TGATGCTGGC GGCCATGTGG TGCAACGGGT
2351 CGGTGTTCGG CATCCAGAAC GCTTGCGGGG TGCTCTTCGT GTCCATGCTG
2401 GAAACCTTCG GCTCCAAAGA CGATGACAAG ATGGTCTTTA AGACAGGTGA
2451 GGCGCGGCGC CCGCCGAGGC CAGCCTGGGC GACCCGCGTG GGGCCCCCGA
2501 GCGCATCCCG CGTGTGGGCT GTGTCTGCCT CCGAGTGTGC ATGTCGGTGG
2551 GTCCCTGTGC CAGAGGGTGC GAGCAGGGGG GTCTTTCGAG TTGCAGACAG
2601 AGCCTGCCGC TTCTGGGGCC TCGGGGTGCC CGTCTTTATA TGGAATCCAG
2651 CTGCAGAGCT GTGTGTTTGC AAGCAGGTCG CAGAACTTAC TTGCCGAGAT
2701 CGTCCTCCTT TCCCCTCAGC AGAGCAGACG CTAACAGTCC ACAGGAGCCC
2751 TTCCTTTTAT TGTTTGAAAA CAAACAGAAC CCCAGAACCT TCAACCCCAG
2801 TCATCGCCCT GTCATTTTTG TGGTCTCTTT CGTGACTATG CCAGTTATGT
2851 AGTTCTTCAC CTGCCTCCCT GGGCCGCAGA GGGGTGTGCG TATGTTGGCG
2901 GGGCGGGGGG TGGAGTTTGG AGGAATGAAA GAGATTTGTA CGAAGGTCAC
2951 TGGAGTTCCA AAGGGGGCCC TGCAAGAGTC ACGGTTCCGT GCGTTCCCGT
3001 CCCCCCGCCT TTTTTTGCCC TCTGGGTTAA ATGTAGAAAA CACGGGAGGC
3051 AGCCGGATTA GGGACTAGGA TGAGGAAGGT GAAGGGTTGC TTCTTCCCTC
3101 TTCCTGTTGT GTTTTTTGAC ATTTTTTTTT AACCATATAG TAAATTAGAT
3151 ACAAAAGGTG CAGATTCAGC GTTTTCTCCC TGTAGAGCAT TATTATGACT
3201 TTTTGGCTGG TTAGGCAAAA AACAAATCTA AGACCTTCTG CATGACACTT
3251 TAACATAAAT TCTTTCACTT TATCCTGCAA GGTGAGCGCG GTCAACCCCA
3301 TTTGGGTGAG AAAACTGTAG CTCAGTGAAA GTGTCTTGGT GGGTAGTAGA
3351 ATGGCAATAA AACACATATC AACTGACTTC AAGGGCTAAG TGATTTCCAT
```

FIGURE 3-1

3401 TACTAAATCA ACCTCCCTCC CCATCATTTG GGGTAACTTT ATATGATTAA
3451 TAGTCTTTTT TTTTAACCTT GATTTTCTAT TATTTTTAGA GTGAATATTT
3501 CTTAGGTCTT TAGTATGCAT ATGAGGAATG GGCAAGACTG TAATAAATTC
3551 TGAGACAAAG GTAATGCTGG GTTATGCTGA GAGTTTTAAA ACCTGACATA
3601 AATACTATTA AACTATTTGT ATCATTCTGC AACTTACTTT TCTTCCATTC
3651 CGCATCATGT TTGTGACTTA TCCACATAAT ACCTCAGTGT GAACTGATAA
3701 CTCAAATTCT TTCCATTTTA ACTTAGGTGG TTTGCATTGT TTGACTATAT
3751 TATACTCTAT GCATTCTCCC TCTGATGGGC ATTTAGATTG CTTCCAAACT
3801 CATTCTAAAC AATGCTGCAA TGAATATTCT TGTACACTCT CGTTATGCAT
3851 GTGAATACGG TACCATTTTA ACCTGGAATT TCTGTTCTTT AAATAGCTAT
3901 TGAAACTGCT GTGGTATGCG GGTCAATGGG CTAGGTACAA AAAGTGTTAA
3951 AAATGTAGTA ACATATCCTT ACCATTTAAG GGAAGTAATC ATTGTAAAGT
4001 TTAGCAGGGG AGATATGCAT ATATAATAGC AAACAAAAAT AGTTTGTTGT
4051 CTTTTCTATA TGAGTATTGG GTGTCAGAGA GAAAAGCCCC AAAAGAAGGC
4101 AGAATTGACA GAGTTAACAT TTAAAGACTA GTTCCAACAT TTACCATATT
4151 CCTGCCTGGG ATTACAGATT TTTTAATGCA GTCAAGATAA CAGCAGTCTT
4201 TTGTTTATCA TTGTTTTTGC AAATTCAGTT AAGTAGATCC TTTGGTGTCT
4251 GTGGGTGGGT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTGAG
4301 AGAGAGAGAG AGCAATTGCC AGAGAGACCA TAGCTTTGCC AGGGATGAGA
4351 ATTTTGCAGT GTCAAAGTCT CTACCTACTA CTTGTCCCCA AAGTTCTAAT
4401 TGGCTACACA ATATCCCAAT ACTGGGTAGC TGAGAGTGAG GGAAGGAACC
4451 TGGTTTTTCT TTTGCACTCT GTGGAACTTT GTGTTTTCCA TTTTGATGAA
4501 TATCTTTTTT CTTTTTACTC AGTTCAGTCT TTGACAACTT TTTCAGTCAT
4551 GTTTGTGTAT GTGTGGGTAT ATATCATATA AACAGTTGCA CAGGTGTGCT
4601 AGTTAAATGT GTGAAGATCT TTGTGTTTCT CTGCCTGACT GCTGTATATC
4651 TATTTATGGT TGTGCCATTG CACAAGGGTG CCCAACTCAG GGGTAAGTGG
4701 GGACTGAAAA CCAGCCTGGG CTCTGGGTGC CCTTTGTCTG ATTCCTACAG
4751 AAGGGCCCTA TGCACTTGTA ATGGCCCTGT ATAACACAGC ATCTAGATTG
4801 AACAATGGCC ATTACTTGGG TGCTAGGTAA TACATATATG ACTGATAGAT
4851 GTTAAGGGCT GAGGAAGAAA ACAGATTTAA ACTTAGTGCT GAAAAAATGG
4901 TTACAAGATA GTCTTTAAGC CAGTTATTGT TGAGATCTCT CTCTCTTCCC
4951 CTGTCCCCTA CCCCTTTCTC TTTCCTTCAG TGCACACACA CACACACAAA
5001 GGTGTTTCAT GAAGTCCCTC ATCTACCACA GTCACTGTTA TTTGAGAATA
5051 TCTGCTTTGA AGTTTGATTG GTCACACTTT TTCACTTTGA TATTCGAATG
5101 CTGAGTCGTC TGTGATCAAG CATATGCAAG CTTCAAATAC ATGCCAAAAA
5151 ATATCTGGAA TTTGTTTAAG CCTTTTATTT TTCAAAGTTT TGGTCTATTT
5201 TCTATTACCG TACTCATGAT GGATAATCCT GGTGTTAGAG TACAGCTAGT
5251 TCTGTCTCCT TGTTTCCATT ACTTCTTTAT AGCAAGTGAC TAGCCTAAGG
5301 ATATACAGGG AGGTGGTGGT GGAATGGAAT CTAGGTCTCC AAATGATGGT
5351 GCGCATTTCT TGAGTACTTT CCTGTGGCTA AGCACTTTAG ATGCGTTCCT
5401 ATTTAAACCT TACCACGATT CTCTGATAGA CTTTGTTAAT ATCTTTCTTT
5451 TCAGATATGG GAACTCAGGC TTACAGAGTT TAAGTAAGAA GTGGAGCCAG
5501 AATTCAACCC CAGGCTTATC TGACTCTAAG AGCTGGGATT TTTATTTTAA
5551 TTATTTATTT ATTTAAAATA TGGAATGCTT CATGAATTTG CATGTCATCC
5601 TTGTTCAGGG GTCACGCTAA TCTTCTCTGT GTCATTCCAA TTTTAGTAGA
5651 TTGTTGTTCA AAGTGCTGCT GAAGCAAGCA CCAGGAGCTG GGTTTTAATC
5701 ATTCATCATA TTGCATTGAC TAGATAACAT TCTGCAAATA CGATGTTTTT
5751 TATGTTGTTG ATTAATTTAA GTGTTAGTGA TTGGTTGAGT GCTCTACCAT
5801 GCATTCTGGG ATTAGAAAGA AGGGTCCCTG TTTCTTGGTC CTACTTTGTG
5851 GTGAATAAAC AATTGCAAAT TATTAATGTC TCAAACTATA TTTCTGAAGT
5901 GTAGAGAGAC TTCCATAGAA GAACAAGATA CTTCCATATG CCGTTCAAGC
5951 AAAAGTCTGG GGTTTCCTTT GAAGAACTTT TAGATTGATC CACAGCAGGA
6001 CAATGTTTCT AGGCAGAACT GAGGAGGAGC CTTTCTTAGG CTCACTTCTC
6051 TTCAGGGCTC TGTTAACTCT TCCCACGCAA TGGATAATCT ACCCAAAATT
6101 TCTCAGGAAA GGGCCTGAAG AAGTTCATTC ACACTAAGGT GTAAGTGAGT
6151 TTACACATCT TACTGTTAAT TCTCTTTATA CAAATGTTTA CCAAGTTATC
6201 TAACACGCTT TGTTTTGGGC TCTGTCCTGG GGACTGGAGA TAATGACTGA
6251 GAGAGAAAAT GTCAGCTGTT TCAAAGTAGC TTAGGATCTG TTGTGGGATA
6301 CAAATTAATA ACAGACCAGA AGTAATAGAA TATTTCCCTG AAGGATTTTC
6351 AATATAACAG GACTCAGTTT TACTATAAAA GGCTGAAATT CTAAGGTCAT
6401 TTCAACAGGT GGTGGGGTTG GGGGTGGGGA AGGCATTTGA CGCCTCTTTC
6451 TCTATGGTTA TAAATCTCAC TTGGTGAAAT TAAGACTTTG GAAAGGGGAA
6501 GTAAGCCAAC TCCAAGTTGG GCAGTAGAAC CAATGAAAAA TGCTGACGGC
6551 ATCACAGTCC CATTATGGTG CCCAGCTGCC AATGACATGG CACTCAGAGG
6601 AGTGTCTCAC ACATACTGCT CTGTCTGAGG GAGCAAGCTA AGCTTGAGTT
6651 GTCTCTTTTT TTGTTGTTTT TTTTTTTTTT TTTTTTTTGA GACAGATTCT
6701 CACTCTGTCG CCCAGGCTGG AGTGCAGTGG CACCATCTCG GCTCACTGCA
6751 ACCACTGCCT CCCGGATGCA AGCAATTCTG CCTCAGCCTT CCGAGTAGCT

FIGURE 3-2

6801 GGACTACCTG CGCTTGCCAC CACACCTGGC TAATTTTTGT ATTTTTAGTA
6851 GAGACAGGGT TTCACCATAT TGGCCAGGCT GGTCTCAAAC TCCTGACCTC
6901 GTGATCCACC TGCCTCGGCT TCCCAAAGTG CTGGGATTAC AGGCATAAGC
6951 CACCGCGCCT GGCCAAGTTG TCTCTTTTTA GTTGAATTTT TACCTGTTCA
7001 CATGTGTATT CTTCTTGCCT AGGTAGAGAG GAATCAGACA CTCTGGGGAA
7051 GAATACAAAG AAATACAATT AAGTGGAACA TTGTTTTTCT TTAGAAAGTG
7101 CAATTTTGGG CTGGGCGCAG TGGCTCATGC CTGTAATCCC AGCCCTTTGG
7151 GAGGCCAAGG CAGGTGGATC ACCTGAGGTC AGGAGTTTGA GACCAGCCTG
7201 GCCAAGATGG TGAAACCCCG TTTCTACTAA AAATACAAAA AATTAGCTGG
7251 GCATGGTGGC GGATGCGTGT AATCCCAGCT ATTCGGGAGG CTGAGGCAGG
7301 AGAATTGCTT GAACCTGGGA GGCAGAGGTT GTAGTGAGCC AAGATGGCGC
7351 CACTGTACTC CAGCATGGGC AACAAGAGTG AAACTCCGTC TCAAAAAAAA
7401 AAAAAAGAAA AGAAAAAAAG AAAAAAGAAA GAGCAACTTT GTTTTAACTC
7451 TGCTAGATAC TGGAAAACCC ATGGAACTAA TGAAGAGCCT AGGGCTTTTT
7501 ATTTGTTTTG AGATTGTGCC ATTTCACTCC AGCCTGGGCA ACAAGAGAGA
7551 AACTTTGTCT CACACACAAA AAAAGTGTAA ATCAAAACAT TAAAAATTAA
7601 GTAGTTTGGA AGTAGATTAT CAAAAAGGTC CTGAAAGGGA GGTTCTTTGG
7651 CTATAATCTT TAACGCAACT CTACACTCCC TGTATGGAGA CAGATTTCTT
7701 TTTAGATGGT TACAGTCACA AAGTAGGGTT TTCAGTAGCA TTTAGGGATG
7751 AATGAATCTT GCAGCACCTC TCCATGTATC TTGCTAGCCC CTCTGAAACT
7801 TCAGGTCAGT TAGTGCTTCC TCAGAAATTG TTCCCCCCAC ACCAAGTTTT
7851 CACATTTACA GTTATACTGA TATCCACATT GTACTGTTGT ATGTGACACC
7901 TAGATTATAG GAAATTTTGG CTATAGTTCA GAAATTAACT GCTATGTTTT
7951 GCCTTTACGC TAAAGAGATT TTGTTTTGTT TAGTAGGAAA AGCGGCCTGC
8001 ATAACTAGCC ATTTCTGTAT CTTAGAAAAA TTTTTAGTAA CAGTCCTTTG
8051 TTGAGCTAGT TACAGTGAAC AAATAATCTG GTTCATGGTC CTATACATCT
8101 TTCACTATAA GAAAAATACC TGATTGTTAT TTACACTGGA AGAGAGGTAG
8151 AAAAGCTAAG AGAACTCACT TATGGCAATA AACCAATCTA AACTACCTGC
8201 TAAAATAAGT GAGAAGATTA TAAAAATGGT TCTAGGATTT TGGAATAATA
8251 GTGAGTATGG TATGGGCGTT TCATACTTCA TTTCCAGAAT GTTTCTGGAT
8301 TAAGTGCGAG ACTGAATAGC ATATATAGTG AATTCTAATT AAATACAACA
8351 ATGTGAGATT CCTGTGGTGT TTTTTCATGG AATTAAAAAT TAATAATTTC
8401 AATAAAATTA ACTGCTGAAA GAACCCAATT AGCCAAAATG AAAAGCATAA
8451 CACATTTTTT CAGGAGCGAT TTTGAGGTGT CTTTTAGAAT AAATTGTACT
8501 CTGCTTTTGA TGTGATTTGC TACATCTTTT TGTTGCAGTT CCTTGAGGCT
8551 CAGCCCCTGG CCATATACTT GCTTCACTTT TCCTGCTTTC TTCCATCCAC
8601 TGTCTTGGGG CTGTTATTTC CAAATCTCAT CACTGTGTTC AAGACTTATT
8651 TACTATTTCT GGACAGTTCC ATTTGGGTAC ACAGGTACAT CATACTAAAC
8701 TAACATGAAC TCATTTTTCA GCTACACCAT ACAACCTTCC CTTACTACCA
8751 AAAATGACAG CCCATTTGTC GGCATTCTTC TGAATCCATA TTCCTCCTTC
8801 TTAATTCTCT GTGCATGATA CCTCTGGTTG TTTAAGTCAG AAACCTGGAA
8851 TGTATCTTAG CTTCTTATTT CTCTTCCTTC CTTACCTGTT TTGTTAAATC
8901 AGAGTTCTTT TGTGTTTACC TTCTTAATGC CTCATAAATC AGTCCCACTT
8951 TTCTTTCACT ACTGTGTCTT AATTCATGCC TTCATTTACT TTTTATTTTA
9001 TTTTTTGGGA TAGGGCCTCA CTGTGTTCCC CAGGCTGGAG TGCAGTGGCA
9051 TGATCATAGC TCACTGCAGC CTCAAACTTC TAAGCTCAAG CAATCCTCCT
9101 ATCTCAGCTT TTTGAGTAGC TGCAACACCA GGCACATGCC ACCATGGCCG
9151 GCTAATTAAA AAAAATTTTT ATGTGGAGAT GGCATCTTGC TATTTTGCTC
9201 AGCCTTGTCT TGAACTCCTG GCCTCAAGCG ACCCTCCTGC CTCAACTTCC
9251 CAAAGTGTTG GGACTACAGG CGGGACCTAC TGTGCCTGGC CACCTTCATT
9301 ACTATTGGCA ACAATTAGTC ATAACCCCTT AACAGGATTG CTTGTCCTCA
9351 GTTGTACATC TGAGTGATTT TTCTAAAAGA TTGGACCATA TGATTTTTCT
9401 GTTTAAATGC CCAGTGACAC TCATTACTTT TAGGAAAATG TCAAACTCCC
9451 TACTCCGAAG GCCTGCAAGC TCTGGCCCTT GCCTGGCCCT CTAGCCTTGC
9501 CCCTGCTTCT CTCCCTTACT GGTCTTTGTG TTCTAGCCAA CCTGTAGGTG
9551 TTACACTGGC CCAAATTTGT CTTGCTGCTT TTTGCCTCTG TACCTTTGTG
9601 TGTGCCACTC CTGTCTTCAG TGCGATGGTT GGTCCTTGTG AGATTCTGAT
9651 GAAATGGTTG GCCATTTTAT TCTTATGTCA CAATCCTGGG ACACAAACAG
9701 TGATTTTATG CAATTGTTAT GTATTTGATG CACTTGGAAT ATTGGGGGTA
9751 GCTACATTTT GGAGTTTTGA GAACGAATTC AAATAAGTTA CAAATTATGT
9801 TTAAAGTGGT AGACAGAGAA CCTGATTTCA ACCTATTCTA ATAAAGCATT
9851 CCGTGAAAGC CATTTTAAAG ATGATCCATA TTTGTTAAAG TGGTAATTTT
9901 TATATTCTCT GATATGGTTT GGCTGTGTTC CCACCCAAAT CTCATCTTGA
9951 ATTGTAGCTC CCATAATCCA CACTTGTCAT GGGAGGGAGG TAATTACCTG
10001 GTGGGAGGTA ATTGAATCAT GGGGGCAGGT TTTCCTGTGC TGTTCTTGTG
10051 ATAGTTAATA AGTCTCATGA GATCTGATGG TTTATAAAGG GCAGTTCCCC
10101 TGCACACTTT CTGTTGCCTG TTGCCATGTA AGACATGCCT TTGCTCCTCT
10151 TTCACCTTCC ACCATGATTG TGAGGCCTCC CTAGCCATGT GGAACTGTGA

FIGURE 3-3

10201 GTCCATTAAA CCTCTTTTTC TTTATAAATT ACCCAGCCTC AGATATTTCT
10251 TCATAACAGT ATGAAAATGG AGTAATACAT TCCATTACCA TAAAGAAAAG
10301 GCTTTCATGT ACATTATTTT TTAGAGTAGC CTTGTGGTAT GTCATTACCT
10351 CCATGGATAG ATAAGAAAGT TGCAACTTGC ACAGTATTAG GATTGATATC
10401 AGTATTTACT TTTATTAAGT TGAACTTAAG AGCAGCTTTT TGGCTGGAAA
10451 AAAGTTGTAC TTATGTCAAA GTTGTCCTGA AAGTAGAATC CTACTCCTGT
10501 CCCCAGCCTG AAACTATTTA CTACATATTT ACTTGCATGT TCTTTAGAAT
10551 ATTCTCTCAA TAGTGTCTCC TACTCAAGTC ATCAGAAATG CTGTGATGTC
10601 ATTTTTGGAA AAGAATTCCA GAGTTATCAC CGCCTAGCTA GAAATCTGGT
10651 CTTATATTCA AATTAAACAA GCAAACCTTA ACAAAAACAA GCTAAACCCT
10701 AAACATAACA TGAACAGTCA GCTCACCACA GTTCTGAGCA CCTGCCTTGG
10751 CCTGGTGCCA CCCACGCAGG GACTGTGGAT GTTTTTATTG GCAGAGATTC
10801 AGAACAGGAA ACTCCAGCAC ACCTGGGAAC TGCGCAGACC CACCACATAA
10851 GACAGATAGC CTATCAGTGG CTGGAGGAAT GGAGGAAAGC AGTGCTTTCA
10901 AATGTACATG CCAAATGTGT ATGATCATAC CTCTTTGTTA AAGTGCCTTC
10951 TTTAACAGCA AAAGTAATTC CTCACCTTGC ATATAGGAAC TAAAAAAAAG
11001 TCGATGAAGA AATGGCTTGC CTTATTTTCA AGTAAGAAGT CTTTTTTCAT
11051 TTCACTAATT TTTAATTATG GGCATAAGTA TGAAATACAG ATTAGAAATA
11101 CTGAATGTGG ACCAAAGCAA TGTTTCCTTT GTGGACCAAA GCAGTGAATC
11151 TTTTTTCTTT CTTTCTTTCT TTTCTTTTTT TTGTTAAGAG ACAGGGTCTT
11201 GCTCTCTTGC TCAGACTGGA GTGCAGTCAG TGATGTGATG GCTCACCATA
11251 ACCTCAACCT CTTGGGCTCA AGGGATCCTC CTGCCCCAAC CTCCTGAACA
11301 GCTGGGATTA CAGGCACATA CCACCACACC TGGCTAATTT TTAAAAATTT
11351 TTTTGTGGGG GAGGGTCTCT CTATATTGCC CAAGCTGGTT TCAAATGCCT
11401 GAGCTCAAGT GATCCTCCAA CCTCAGGCTC CCAGAGTGTT GGGATTACAG
11451 GTTTGAGCCA CTGTGCCTGG CCCCAAAACA GTGTTTGTAA CTCCCTTTGT
11501 CCCCTCTTGA TAAACATAAA AGTCTCATGG TACTTCAGAA TTTTGTTGCA
11551 CACATCCAAG TGTAGTTTGC CTTTCCTTGT GAGTGGCAGA AGACAATGTC
11601 ATACTCTGTA TTTATCCATC AGCCAAAATT TTGTCAAGCT TTACTTTTAT
11651 TTTTTAAATT TTTATTTGTA TTTTTTTTTG GAGACAGAGT CTTGCTCTGT
11701 CGCCCAGGCT GGAGTGCAGT GGCGTGATCT CGGCTCACTG CAACCTCCGA
11751 TGCCTGGGTT CAAGCGATTT TCGTGTCCCA GCCTCCAGAG CAGCTGGGAC
11801 TACAGGCACG CACCACCATG CCTGGCTAAT TTTTTTGTAT TTTTAGTAGA
11851 GGTAGGGTTT CGCCATGTTG GTCAGGCTGG TCTCGAACTC CTAATCTCAG
11901 GTGACCCACC CGCCTCAGTC TCCCAAAGTG CTGGGATTAC AGGCATGAGC
11951 CACCATGCCC GGCCAAGCTT TACTTTTAGT TATGTTGTGG ATATAATTAG
12001 AATTATTTTC TTGTTCTTTA AATACAGAAT TATAATATTG AATGTGTGCT
12051 TTAAAAAAAT TAGTAAATGT ACCCCTCTAG AACTTCTGAT CTATGCCTCT
12101 TAGTGAGTGA GGGGAGCTGT GCCGCTTTCC TCCTCCAGTG CCTCCACTTA
12151 AGAATCACTC ACTAAGGAGT TTTAAATTCA ATTAAAGGTA TCCTTTAGAT
12201 AGTTAAGTCT AAATGAAAGG TCAATGATTA AATTAATGGA TAAAAGTCCA
12251 TTGCACCTAC GGAAAGGTGC ATTGGTCTAC AGTTCAGCTA GTACCTATTT
12301 TTTGCTAAAG ATTAAATTGC AGCTGGGCGG GGTGGCTCAT ACCTGTAATC
12351 CCAGCACTTT GAAAACCTGA TGGAGGCCGG GTGCCGTGGC TCATGCCTGT
12401 AAATCTTAGC ACTTTGGGAG GCCGAGGTGG GCGGATCACT TGAGGTCCGG
12451 AGTTAGAGAC CAGCCTGACC AACATGGAGA AACCCTGTCT CTACTAAAAA
12501 TACTAAATTA GCCGGGCATG GTGGCGCATG CCTGTAATCT CAGCTACTGG
12551 GGAAGCTGAG GTAGGAGAAT CGCTTGAACC CGCGAGTTGG AGGTTGCAAT
12601 GAGCCAAGAT CATGCCATTG TACTCCAGCC TGGGCAACAA GAGTGAAACT
12651 CAGTCTCAAA AAAAAAGAAA ACCCTAGGTG GAAGGATCGC TTGAGTCCAG
12701 GAGTTCAAGA ACAGCCTGGG CAACATAGTG AGACCCCATC TCTACTTTTA
12751 ATTAAAATAA AATTAACATT AAAAAGTGTA GTAAAATTTT TAAAAAGAGT
12801 AAATTGTATC AGCAGTGTTC TGCCTGTGTT CAGAAAGCCA AAATTTATAT
12851 TTGTGTTTCA TTTAGATTGG ATCCAAGACC AATTTTGAGA TGTGTTTTAA
12901 TATTACAAAA ATAGAAAACT ACCTGTTCAT TAAATGGTGA TATTTATTGA
12951 CTTTTCTGTG TAGATAAGTA TTAATGCCAA GTCAAGTAGG TTAGTCTGGA
13001 GACTGTTTTT ATTAATAAAA ACCTTTTCAT CTTAATTATC CTTTTTATTA
13051 GTTTTGACTT TATGTTGCAA CTTCAAAGCA GCATCTCAAA GGGTTCACAT
13101 TTGTCAGTGG TTGTTTAAAC AACGACATTC CACAAAGTTA ATTGATGTTT
13151 TAGTGTGAAT GGGGCAGGAA GTTGTCATTG TTGTCCTGTA GTAACGTAGC
13201 ACTCAAACTT TGACAGGAGG ACCCCAGCCA TGTAGTTATT GACTGTAAGT
13251 ACACTTCCCT TAAGAGTAAG TTGTTTTTTA GGAATAAATG ATAAATGTAT
13301 GGAGTTTTAG TTGTAGGTGT TATGTTTTGT CTCCTTCTTC TCAGAGAAAA
13351 ATGTTACTGT GGAATAGGCT TAAACTGAAA AAAGGATCTG ATTTTTAATA
13401 AGATCAGATT CTTTGGCCAT ATTTTGATAT TGGTTCAAAA CAAATGTTTA
13451 ATATCAGATG CACAATGTTA AGAGCTCTAT AATGTAGTGG TAACACCTGA
13501 GCCTCAGCCA GTGACTACAC ATTAAGTTCT TGTTCATTTT TTGCAGGGGC
13551 AGGGGAGCTG GTGAGGCATA AAGAAGGGTC AGAGGAGATA ATAGACTTAT

FIGURE 3-4

13601 ATTTGTATTT TTATGCATAA TTATATAGGA AACATTAAAT CCAGAGTTGA
13651 TAAATAATTT GTATGTATGT ATTTATTTAT TTTTGAGACT GGGTCCTGCT
13701 CTGCTGCCCA GGCTGAAGCG TAGGGGTATG AACACAGCTC ACTGCAGCCT
13751 TGACCTGGGC TCAAGCGATT TTCTTGCCTC AGCCTCCCGA GTAGCTGGGA
13801 CCACAGGCAT GTGCCACTAC ACCTGGCTAA TTTTTAAAGT TTTTTTGTAG
13851 AGATGGGGTC TCACCACGTT GCCCAGGTTG GTCTTGAACT CCTGGGCTCA
13901 AGCAATCCTC CTGCCTTGGC CTCCCAGAGT GTTGGGATTA TAGATGTGAG
13951 CCACCATGTT CAGCTGATAA ATAATTTCTA ATCTAAAAAT CCTATTTTGT
14001 ATGGAGAGGG GAGGGCAAAT AGGCTATTTT TTCCACATTT TGTTGCTGGC
14051 CAGAATCTCA GAGGGTTTTT ACCTGCATTA AAAATGATTA AGGCTGGGCA
14101 GAATGGCTCA TGCCTGTAAT CCCATCACTT TGGGAGGCTG AGGCAGATGA
14151 ATTGCTTGAG CCCAGAAGTT CTAGAGCAGC CTGGGCAACA TGGTGAAACC
14201 CCATCTCTAC GAAAAATGCA AAAATTAGCT GGACGTGGTG GCAGGCACCT
14251 GTAGTCCCAG CTACTCAAAA GGCTGAGGTG GGAGGATCAC CTAGCTCTAG
14301 AGGTCAAGGG TGCAATGAGC CAAGATTGCT CCACTGCCCG CCAGCCTGGG
14351 CAGCAGAGCA AGACCCTGTC TCAAAAAAAA AAAAAAAAAA GGTTATTTTT
14401 TTTCAGCATA GACAAAGCAG GGGAAGGAAA AGATTATGTT TCAAATGTTC
14451 ATTTAAACTA CTACATTTAA AGTAATACTT CCTAATGATT TAAACTTTAG
14501 ATTAGTCTAT TTATGGGTCA CCTGGAAGAT TCTTTATAAA ACATGAGAGT
14551 TTATTACTTC TTTCAATACA CGGGTGTCTG TAAATGATGC TCAATAGATC
14601 TGAAGCCTGA ACTTTCTGAA GAAATGTTGT GAAATTATCT ATGGATGTCT
14651 ACTTGAATTT AAATAAAAAT AAATTGTAAT ACATTTGGTT TTATTGGTTT
14701 TGAATTTGTA ATTATTTGGG GATTTGGATT TGGTTTATCT ACAGTTGTCT
14751 TTTTTTTTGA GGTGGAGTCT CGCTTTGTCA CCAGGCTGGA GTGCAGTGGT
14801 GCGATCTCGG CTCACTGCAA TCTCTGCCTC CCGTGTACAA GCGATTCTCC
14851 TGCTTCAGCT TCCTGAGTAG CTAGGATTAC AGGCATGCAC CTCCATCCCC
14901 AGCTAATTTT TGTATTTTTA GTAGAGATGG GGTTTCACCA TGTTAGCCAG
14951 GATGATCTCG ATATCTTGAC TTCGTGATCC GCCTCGGCCT ACCAAAGTGC
15001 TGTGATTACA GGTGTGAGCC ACCGTGCCCG GCCTACAGTT GTCTTTTTTT
15051 ACTCACTCCC ACAGATGAAT CATTATAAGG AGGTTAGCTT TCCTTAAAGA
15101 ATACACTCTC CTTAAGCGGT TTTATCAACA AAGCCAGGGA ATGCCAAACT
15151 TTAACACTTT TACCTAAATT TATAACTGAT GCATGTATGC ATATATACAT
15201 ACATACATAC ATGTATATGT TGTTATATAT GTATGGGTAT CATCAGTATA
15251 GTCTATCAGT ATAGTAATTG TTTATCTGAA ACTTGGGGTT CTCTCTCGCT
15301 CTTTCTCTCC CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTATATAT
15351 ATATATATAT ATATATATAA ATATATGTAT ATATATTTTT TCTTTTTTTG
15401 AGACAGGATC TCATTCTGTC ACCCAGGCTG GAGTGCAGTG GTGGGATCAT
15451 GGCTCACTGC AGCCTCGACC TCCTGGGCTC AAGTGATCCT CCCACCTCAA
15501 CCTCCCAAGT AGTTAGGACT ACAGGGGCAT GCCGCTACAC GTAACTAATT
15551 TTTGTATTTT TTTGTAGAGA CAGGGTTTTG CCTTGTTGCC CAGGCTGGTC
15601 CTGAAATCCT TGGCTCAAGC AATCTGTCCA CCTCAGCCTC TGAAAGTGCT
15651 GGCATTACAG GTCTGAGCCA CTGCGCCCAG CCTAGATTTT TTTGAATTGT
15701 AAAAAAGTAA CCTGCTCCCT ACTGAAGTAA ATAGAGTTAA AAAAAGTAAT
15751 CTGGTACAGA CACCTGTATT TTCTGACACC CCTAGAAGAG TCCCAGGTAC
15801 CCTATAATCA AATACATTAA CATTTCTGCA GCAAAATGTA TGGATAAGTG
15851 AGTTAAATAG AGACCATGAG TAGCTTCAGG TCAGTTCAGA TCAAGTTTTG
15901 CTTCTAATTA AATGTTGATA TTCTCTTACA AAAACTTTGG GTTTGGGTTT
15951 TCAGATTTGC AAATAAATAA TTATAAATTA TTATTTTTTT TGAGACAGAG
16001 TCTTGCTGTG TTGCTCAGGC TGGAGTGCCA TGGCACGATC ACGGCTCACT
16051 GCAACCTCAA CCTCAGGCTG AAGCCATCCT CCCACTTCAG CCTCCCAAGT
16101 AGCTGGGACT ACAGGAGTGT GCCAACATGT CCAGCTCATT TTTGTATTCT
16151 TAGTAGAGAC AGGGTTTCGC CGTGTTGGCC AGGCTGGTCT CAAACTCCTG
16201 GTCTCAAGTG ATCCGCCTGC CTTGGCCTTC CAAAGTGTTG AGATTATAGG
16251 TGTCAGCCAC TGGGCCTGGC AGAATTATAC ATTTATATGT CAATATTTGC
16301 TTTTGTTTTC TGTTTTTCAG TAAACGTTTT TTAAGGTACA TTTTTCTGTA
16351 TCTCATAAGG CACCTGCTTA ATTGTTTCAG TAAGTGTGAT GTTCTACCAT
16401 ATTGGTCTAC CCTAGGTTAC TCAACCAGGC CTCCTTTGTT TAGTGAGTAG
16451 CAGGCAGTGT TGTACAACAT ATGTAGCATA TCTGTATATG TCGTCGAACA
16501 AATTGTTTTT TTCCCCTCTC TTGGATTGCT TCCTTGGGTG TACGTCCAGA
16551 AGTGAGATTA CTGGTTCAAA GGGTATGAAC AACTTTATAA CACCTGTTAC
16601 ACATTGCCAG ATTATTCTTT AGAAAACTTG AATCAGTTTA TTGTGCCACC
16651 AGTGATGTGC TGGCTTCCTG AAAACCCTAC CAATGTTTGG TTTTATTTTT
16701 ATTAGTATTT GCTAATTTGA TAAGTACTAA TGATATTTTT TAAAAGTAGT
16751 TTAAAATCAT ATTTCAGTGC TTATAAGTCT GTGTTCCCAG TTTTTGAGCC
16801 CTTTAGAAGC TGCAAATGAC CTGGCAATTA TATATAATAT TTGAAAATAC
16851 AAGAGGACAT ATGCCAGTGA ATATATTAGA GTAAAACTTC ATTCCCATAG
16901 GTAATGAAGG AATGCTTGAG ATTATCTTAG GCCTTAGATT CTCACCTGAC
16951 ACATCTTGGC AGGTAGACCA TGTCCTTGTT TCCTCTGCTG TCTTAGCCCA

FIGURE 3-5

17001 GGTGTTGATC AAGGTCTGTC TTAGGGCGGG GGATAGGAAT GGAAATAAAC
17051 CATGTAGAGA CTTGGGCATG AGGACTTTGT GATTCTTCCA GGTGACATCT
17101 CATCCTTCAG AGGATCAAGT CTGCAAGAGT AGCCATATCT TAATCTCTTT
17151 CAGTGCTATC ACCTTGCATC AACCTCTGGA CTCGAGCTAA TTCCGTTGAA
17201 AATATTTTAT TAATTAATTT TGGGGTATGT TAAAAATTTT GTTTGCATGT
17251 ATTTATTTAG TTATTTTTAT GAGACAGGGT CTCGCTCTGT CACCCATGCT
17301 GGAGTACATA CGGTTGCACG CTCATGGCTC ACTGCAGCCT TGACTTCCCA
17351 GGCTCAAGTG ATCCTCCCAC CTCAGCCTCC TGAGTAGCTG GGACTACAAG
17401 TGCATGTCAC CACATTTGGC TAATTTTCAT ATTTTTTGTA GAGACGGGGT
17451 TTCGCCACAT TGCCCAGGCT GGTCTCAAAC TCCTGGACTC AAGGGATCTG
17501 CCTGCCTCAG CCTCCCGAAG TGCTGAGATT ATAGGTGTGA ACCACCGGGC
17551 CCGGCTCCCC ATTTAATTTT GTTGTTGTTG TTTTTTAGAT GGAGTTTCAC
17601 TCTTGTCGCC CAAGCTGGAA TGCAATGGCA CGATCTCGGC TTACGGCCAC
17651 TTCCACCTCC TGGGTTCAAG CGATTCTCCT GTCTCAGCCT CCCGAGTAGC
17701 TGGGATTACA GACACACGCC ACTATACCTG GCTAATTTTT GTATTTTTAG
17751 TAGAAATGGG GTTTCACCAT GTTGGTCAGG TGGGTCTCAA ACTATTGACC
17801 TCAGGTGATC CACCTGCCTC GGCCTCCCAA AGTGCTGGGA TTATAGATGT
17851 GAGCCACCAT GTCCAGCCAC CCATTTAATT TTTTGAGCAC AAAATATGTA
17901 CTGAGAGCCA CGCAAGAAAC AAATTCGACT TATTCCATGC TCTTGAGAGG
17951 TCATGAGGGG AAACAAAATG ATACATAAGT AACTCTGAGA GAATATGCTG
18001 CACATGCTAA ATCCTGTGCA AGTAAGATAT AGGATTTTAG AGGAAGGGAG
18051 AATGACTTCT GATTGAGCTG ACTAGAGAAG GCTTCAGTTT TTGAGTTAGG
18101 TGTTACGAGA TTGGGAGACT TTTCTCAGCA TATCTAACAG AAGAGGGTAT
18151 CCGAGGTGAG AGTGTAAGGC CTGGGCAAGG GTTGGGAGGC AGTTCTAATA
18201 CTGAATGTTC TGACTGTGGT TTACTATGTA TTTCAGGTTA TTTTGTTTAA
18251 TCTATCCAGT AATCCTTTCA TGTAACAATT ATGATGTGTG TGTTTTAGGT
18301 GGGGCTACTA AGGCTAGTAA GTAGTGAGGC TGGATTTAAA CTTAAGTCTC
18351 CAGCTTCGTG GCCCAGGTTC TTTATACTTG ACTCCACACT GGGCTTATTA
18401 AGTGAATGAC AAGGAGTTTG ATTTGTCGAG GGCCGAGGAT ATGTAGTGAA
18451 AGGTCATAGA CTATAAGGCT GGCAAAGAAT GGTGGAGCTG GGGAACGGAA
18501 GATCTAAATG CCAAGGTAAG CTTGGACATT ATTTAATAGA GGCAATGGGG
18551 AGATACTGAG GATTTCTGAA GAGGGTTTCT AAGTAGTGCT TTAAAAAGAT
18601 CTCTGGGTTG CAGAGGCAAG AAATAGGAGG AACACCAATT ACAAAGCCCC
18651 GACCGAAAGC CGAGTGAGAT TTGAGGGTCA GCCATGGAAA TGGGAATGAA
18701 GGGACAGATG TGAGACATTT CAAAGGATAG ACTGCAGGAA TTCTGGCAAA
18751 GAATGGGTTG GACAGTGATG AAGAACACAG AAACAGCATT GATGCCTACA
18801 TCTGGAGCCT GTGTCTTAAT AATTTTTGAT CAGGGAGACC TACAATACAT
18851 TTTGTATTGT GCCCCACAAA ATCATGTCGA CAGCCCTAGA TGATGTCATG
18901 GACAGGTATG GGGCATGAGG AGAAAGAGTT GCCAGGGAAA GGCTGGCTCA
18951 CATTTTAGAC ATGTGGAGCG TTAAGGGTCA GCAAGACATC AAGGGCGATG
19001 TCAGGCAGGT AACTGGAAAT GCAAACTGGT AGTCTCAGAG TGACGCTGAG
19051 TCCAGTGGTG GGGCTTTGGA AAGAACATGT ATCTCTAAGA GGAACATGAC
19101 ATAATGGAAA CCCAGAAATC CTGTCTTGAG CGGACTCAGG GGCTGCTTCT
19151 AGGTAATTTA GTTCATTTCT ACTGAAATCA TTATATTTAA AAGTATGGCC
19201 GCACTGAGAT GGCCACTGTA GCTGCTGCCA CCTCTTAGCT TTGGTCTTAA
19251 AAAAAAAAAA AAGTAATAGA ACTTCCTTAA AATGTCTTTC TAGCCTTTGG
19301 ATTTCTCTAA TTCAGATTTG CGTCTTCCCA AGGGTCAAAA TTATATTTTT
19351 ACTATCCCTG TCTTAGGTAT TTCCAAAACT TCGTCTTAAG ACTTAGTCAT
19401 TTTTTTCCTT CAATTTGACA TGACTGCTAA AGACTTTTGG CATGTTCCTC
19451 CTCCTTTCAT TTGTGATGTA ATTAAGTTGG TCTGTAAGTC TTATTTTTAA
19501 GATGTTCTAG ACCAAGAGAC TGTGAGAATA GCTTACAGTC ATTTCAACTA
19551 ATTTATGTAT TTTAAATTTA AAGTATTGAC AGTGGTGAAA ACCTGTTCAA
19601 CAAGCAGATG ATGTTATCTT ATATATTCAC AGAGTTTAGT AACTGAGCCA
19651 ACTACTTCAT TCACAGTTCA AAATGAAAAC AGCTAATTCT TTTAAGTAAG
19701 TATAGATTCT ACTCTTTAAA AGAGTTATCT AGGAGAGCTA CTATAATACT
19751 ATTATAGAAT AAGTGAAATT AAGTTATCTG ACTGGGACTG GGGTGGAATA
19801 TTGCACAGTT TATTGAGAAA TATGGACTGT TTTTGCACAT TCATAATGGA
19851 CATTTGAGGT TTGTTAGGGG AGGAGGTGTC ATCTTTATGG CACTTTCTGG
19901 CTGGGAAGGG AGTCAGTCCT AATTGAGATA ATAACTAGCC ACCTGGCCAC
19951 ACACAAGTGT GTTTTGCCTT AGTTACCTGT CACACACTGA GCAGTGAGAC
20001 TCAAGAGAGT GTCAAAGTAC TTTTCAATGC ATAAAGCACT ACAGATCTGT
20051 CCACACTGTT GTGAGTGAGC AGGTTGGCAC GGTGCCTGTG TGCGGGCGTG
20101 TGTGTTGACT CACGTGCTGT CCTGTGATCT CTCAGGACTC AGGTCCTGAA
20151 TTGCTCTGTT GTGACTGAAG CCCAGCTGAA GGTGCTGGAA GTGCAGTGAC
20201 CCTGGAGGAA GAACCAGTAA CAGCAGAGGG TGGATGAAAG GGAATTGATA
20251 GTTGTGTAAG AATAGATATC TGCCGTTTTT TGTAAGCCAA GACACCTTTA
20301 CCCTTCCAGT AATTGTTTCA TCTTTTAATA TCATTTGGCT TCATTTACAG
20351 AATCTGTATT AAAGCAAAAG TCAGCATGTA AGGTGGTATT TTGACCACAT

FIGURE 3-6

20401 TTGTCTGCTG TTGTGCCTTC TGGGTGAACT GAAGTACTGG CTTACTGACT
20451 AGTAAATATG TTTTCTGACA ATTATAGGGA AGGGAAGAAA AGGAAAGTCC
20501 AATTAAAGCA TTTTTCTCCT CAGAGTTTGA AAAATAGAAT TCATGCAATC
20551 TTTTAAATTC CATGCCAACA CATCAGACAA GAAGAGACTT GATAGTAGTA
20601 AAGGTTGGGA ATCAAAAGAA CAATGTAAGT TTTTGATATT GACTTCAAAA
20651 CATGGTGTTC TATAATTTAG TGTTCATTTG TTACGTGTAT GGTATTATAA
20701 TTAATTTTGT ATATGTGGTA GTTATTTTTT GTACTTTGAT TAGGAAACAA
20751 ATATTGAGCC ACTTCAAGAG GCAGACTATT TTGGAAAAAA AAGTCTGGTA
20801 AAAGTAATGC TTAATCTAAT TAATCTGTCT TCATCCTCTT AATTCATCAA
20851 GATGACTTTG GGTGTCTGGT GGCAACTGAG AATGGGTTAT GGAAGAACGT
20901 CAAGGCAATG TAATCCCTAT TATTTACGGT TACTTGAGAG GATAAATTAA
20951 TCAGCGGTCA CTAATCTTTG GATAATCACT CTATTGAGCT GGAACTATCC
21001 TTAGTATTTT GGAAAGCAAG TCAGTGAGTT AGAACTGTCA AAACTGATCA
21051 GCTTTTCTAA GCTTAATGAT AAGTGAATAG AAACTAGTTG CCTTCAACCC
21101 TTTCCTCCCT GCATTGCAGC ATGATCATTC TGTAACTCTG GAAATGGTTT
21151 ATGGAACAAC AGTGAAAATA CATTGATACA CTGTCTTGTG GTAGATTTTC
21201 AGATAGGCTT TAGACAAAGT TCAGAGCCTT TCCTCTAGCT GGGGATTAAC
21251 AAAGCTGCCT TCATAGTTAA ATGTTTGCAC CCTGTGTATG CATTTTCAGT
21301 TACTAGAATT AGGTAAGTTA GTGTTATAAA ATTGGTTTGA GTGTGGATTG
21351 TTTAGGAAGT GAGTCTTTTG GTGGCAGCAA TTCTGTTATG CATTAAATAG
21401 ATACATATTT TGAACTAGTC GACATTGTTT CAGTCTGTAT TTATTAGATG
21451 CTGGGGTGGG TATGGGAATA AAGAAACGTA TGAGGGGTCT TGGAAAAGTT
21501 CATGGAAAAA ATGTACACTA TGAAAAAAAA CTGTGCATGG ATTTCAAAAT
21551 ATTTTTGAAC CAAAATCAAC TTGTACTAAC TTGTTACAAC ATGTCTGAAC
21601 TTGTTACAAC ATGTCTGAAC AGTTCGAGAC ATTAAGAAAT GATATCGCAC
21651 CAGTTTTTAA AAAGCGCCTA TCAGGGTAAC ATGAATTCTG CTAAAATTGA
21701 AGCAAGAACA AACATCAAAT TTATGGTGGT GCTTGGGTAG AAGGATGGTG
21751 AAATCATTGA TGCTTTACAA AAAGTTTATG GGGACAATAC TCTAAAGGAA
21801 CCAGCAGTTT ACAAATGGCT AACATCCTTT AAGAAGGGAC GAGATGATGT
21851 TGAAGAGGAA GCCCACAGCA GTAGACCATC CGTGTCAATT TTCAAGGAAA
21901 AAATTAATCT TGTTCATGCT GTAATTGAAG AGGGAACTTT ACATGAAATT
21951 TTAAACAAGT GGGATTGAGA TCCTGTGGCA TATGTCTGAA GGATTGTAAT
22001 AGGAGAGGAA ACATGGCTTT ACCAGTATGA TGCTGAAGAC AAAGCACAAC
22051 CAAAGCAATG GCTACCAAGA GGTGGAAGTG ATCTAGTTAA AGCAAAAGCA
22101 GACTAGTCAA GAGCAAAGGT CATGATAAGA GACTTTTGGG ATGCTCAAGG
22151 TATTTTGCTT GTTGACTTTC TGGGGAGCCA AAGAATGATA ATATTTGCTT
22201 ATTGTGTGTG TTTTGAGAAA ATTAGCCAAA GTTTTAGCAA AACAAACAAA
22251 TACCCAGGGA AGCTTTACCA GAGAGTCCTT CTCCACCAGG ACAATGTTCC
22301 CGGTCATCCT CTCATCAAAA AAGGGCAATT TTGCAAGAGT TTTGATGGGA
22351 AATTATTAGG CATTCACTTA CAGTCTTTTT TTTTTTTTTA GTTGGAGTCT
22401 TACTTTGTCT CCCAGGCTGG AGTGCAGTGG TGCAATCTTG GCTCACTGCA
22451 ACCTCCACCT GCCAGGTTCA AGCAATTCTC CTGCCTCAGC CTCCCAAGTA
22501 GCTGGGATTA CAGGCGTGAT CCACCATGTC CAGCTATTTT TTGTATTTTT
22551 AGTAGAGATG GGGTTTCACC ATGTTGACCA GGCTGGTCTT GAATTCCTGA
22601 CCTCAGGTGA TTCACCTGTC TCGGCCTTCC AAAGTGCTGG GATTACAGGC
22651 ATGAGCCACT GTGCCTGGCC TATCTTACAG TCTTGATTTG GCTTTATCTG
22701 ACTTCTTTTT CTTTCCTAAT CTTAAAAAAT ATTTAAAGGG CACCTATTTT
22751 TCTTCAGTTA ATAATGTAAA AAGGACTGCA TTGACATGAT TAAATTCCTG
22801 GGACCCTCAA TTCTTTAGAG ATGGACTAAT GGCTGGTATC AACTCACAAA
22851 AGTATCTTGA ACTTGATGGA GCTTATGTTG AGAAATGAAG TGTATATTTT
22901 CATTATCTTT TAATTTCATT CTTTAGTGAA TTTTTTGAGG TCCCCTTGTA
22951 TACATTTTAA TCCTAAGGGA ATAAAGAAAG GAGGAAGTCC TAGCCCTGTG
23001 CTGTCTGCCT AGGTACAGTG TCTGAAACAC AGACCAGTAT TCACCCTTTG
23051 AAATTTGAGG TTTCCATTCA GGAGGTTCTC AAAGAGAATA AATGAGATTG
23101 CTATGCAGGT GGAATCAAAG AGCACACGGC TTATTTATCA TAATCAAAAT
23151 AATGCCATTT TCATAACAAA CTTCACCTGC TTATGTACAT TGTAAATTGT
23201 TGCCTTGATA AGCTTCCCGG AGATAAAGTA ATTCAGCTAA GTATTATTTC
23251 CAATCATAAT TTTGTGTCAT TATGAGCAAC ACAATACTAT ATATGGGATT
23301 GATTCACTGC AGAACTGGAA TAAATATAAA TTAGATCTTT AGAAAAGAAA
23351 CGTAGATTTA AAAATCTTAT GTTAGAAGGC TCAATTAATT AAATGTAATT
23401 AATTTTTTAA AATCAGCTTT ATTGAGGGAT GACTTAGATA TTATATAATT
23451 CACAAATTTT AAGTGTACAG TTTGATAGTT CTGACAATCA AACTGTATAC
23501 AATCATGTAA CCACCATCAC AATCATAATA TAGTGTGTCC ATCACCCCAG
23551 GGTGTACCCT CGTGATCCTT TTTGCAGTTA GTCTTTTTCC CTTACATTCT
23601 GGCTCCTGAA AACTTGATCT GCTTTCTGTC ACTATAGCTG TGCCTTTTCT
23651 AAAATTTTAT ATGAATGGAA TCATACAGTG TGTTTTCTTT TGTATCTGTT
23701 TTTCACTCAG CATGATGCTT TTGAGATTTC TCCTTGTTGT GGTATGTATT
23751 AGTAGTTCTT TCTTTTTTAT TACTAAGTAG TATTCCATTG TATGCCTATG

FIGURE 3-7

23801 CCACATTTTT TTTTTTTTTT TTCGAGACAG AGTTTTGCTC TGACATCCAG
23851 GCTGGAGTGC AGTGGTGTGA TCATGGCTCA CTGCAGCCTT GACTTCCCAG
23901 ACTGAGGTGA TCCACCTGCC TCAGCCACCT GAGTAACTGG GACCACAGGT
23951 GTGTGCTAGT CTGTCTAATT TTTAAATTGT TTGTAGAGAT GGGGGTCTCT
24001 GTATATTGCC CAGGCTGGTC TCAAACTCCT GGCCTCAAGC AATCCTTCTG
24051 CCTTGGCCCC TCAAAGTGTT GGGGTTACAA GTGTGAGCCA TCACACCTGG
24101 CCTACCACAA TTTTTTATCG ATTCACATAT TGATGGATAT TGGGTTGTTT
24151 TCAGTTGTTG CCTATTATGA ATAGAACTGC TATGAACATT TGTATGCAAA
24201 CCTTTGTTGG GATGTATGTT TTTATTTCTC TTTTGTACAT TAAATTTAAA
24251 TTTAAATTTT GTTCTGTATT ATTTGTATTT TTAAATTTCT CAAGTGGGTA
24301 ATACTGTGTA CTTTTTTTTT GAAATTAAAA AAATTGTGGC TGAACAAGGA
24351 GATAAAAAAG TAGGAGTGAG AGGACTCTGG AGAGTTACAG GGCTTTGGTT
24401 TAGAGGATTG GATGAATAGT GGTGCTGCCA ATAAAGAAAT TTAAATATGG
24451 CTGATATTTC CTATATTTAA GAAAGACCAA AGAGGGTCCA TTGAAATGAG
24501 TCAGTGGGAA ATCTCTGATG ACTTCAGCCA GCACGCTTTC ATCGGCCTGG
24551 ATATATGGGA AGTGAAGTCT GATTATAGTC TGTGGAGCAG TGAATGGGAG
24601 GAAGAGATAG GGGTACAGGC TAAGAAGGGA GGAAGTCAAG TCAAAGGGAG
24651 AAGTAGGGTG GTAGCTAGAG GAAGATTAGA GTCAAGCGAG GGTAACAATT
24701 TTTTTTTTTT TTGAAGATAG GAGTAGCTTG AGAACTAACT TAAAGAAGGA
24751 GCCTGTAGAG AGGGAGGAGG TGAAGTTACT AAAGGTCTAA TTGATGGGGT
24801 AAGGTTCATG AGCAGATCAG ATCTTTACAA GGAAGGTTCT TGCTGGGGGG
24851 CAAGATTCAA AACCCCTATT CAGAACCAGG AGAGAAGAAA GTAAGAATGG
24901 GAGCAAATGT AGGTAGGTTT GGTGAGGATC AGGAAATGGA GGGGAAGAGG
24951 TCATTAAATG TGGTCCTGGG GTTGAGCAGC AGATTGGAAG AGAATGGCAA
25001 AAGTTTGGGA GCTGATTAGT GATAAGGAAA AGGTTTGAGA ATCCGATGAA
25051 GATTAGAAAC CATGCATTTG TAGTGAAACT TGTTTCTAAG ATTGTGCCCT
25101 TACCCACCTC CAGCTGTGCT CTGATAGGTG AACTATACAA TTGATGTAAG
25151 GCTGGCAGAT AATCAAAGCC AAGAAATTTC ATGTTTTCTA TCTATTTTCA
25201 CTTCCGTGCC ATATAGCTTC TCTCATATAG TACTCATATG TTTTGAGTTT
25251 TTGGTGAAGC TGATTAAACA TTTAACAACT TTTTTTTTTA CATTTTTTAT
25301 TGTGACAAAA CATATATAGC ATAAAATTAC CATTCTGGTA AATTTACCAA
25351 AACCCATTTA TAGCATAAAA TTTACCATTT TGGTCAACTT ACTGTAAATT
25401 ACTGCCAGTT CTGTACAGCC ATTTGTTTAA TGCAGTCTGC TCTAATACCC
25451 TCCAAGGCAA AATTGCAAAT AAAAACCTGT TTTGATTTCT ACTTAATTTT
25501 GGTAAATTTA CCAAAAGGGC AAATTTTATG CTATAAATTT TTTTTGTTTT
25551 GACATTTAAA ATATGTTTTG TAAAAAAAAT TGACAGATAA AATTGTATAT
25601 TTAACATGCA TAATGTGATG TTTTGTATTA TGTATAACAT GATGTATATG
25651 TCATACATAA CATATATACA TTGTAGAATT GTTAAGTCTA GGTAATTAAC
25701 AAATGCATTA CCTCACACAG TTATCATTTT TGTGGTGAGA ACATTTTAAC
25751 ATCCACTCTC TTTAAATTTT TCAAGAATAA AATTTTATCA TCTGGTCATG
25801 GTGGCTCACG CCTGTAATCA CAACACTTTG GGAGGGCAAG GCGGGAGGAT
25851 TGATTGAGAC TAGAAGTTCT AGACCAGTCT AGACAACATA CTGATACCCC
25901 GTCTCTACAA AAAATAATAA AAAAGAATGC AATATTGTCT AAATTTCTTG
25951 AATTTATTTT TCCTACCTAA TTGTCGTTAT ATATCCTTTA ACCAACGTCT
26001 GCCCATTTCC CTCTCCTCTC TAACCAGTCC AGCCTCTGAG ACCATTGTAC
26051 TTTCTATTTC TATGAGATCA GCCTTTTAAG ATTCTACATG TCAGTGAGGT
26101 CGTGTAAGTA TTTGTCTTTC TGTACCTGGC TTGTTTTACC ATTGTCACCA
26151 TTTATAAAGT GTACAATTTA GTGGCATTAA GTACATACAG AATGCTGTTC
26201 ACCCATTACT ACTGTCCTAC AATTTTAATT TGAGGCATGA CTTTGTCATC
26251 TCTCTTCTAT CCTTTGTACC ATTCTTGATC TATAAACTCA ACAAAGTCCT
26301 TAACTGAGTA GAATATTTTT GTATGTGTGT ATAGATGTGT GTGTGTGTAT
26351 ATAGGTATGT ATAGATATAG ATATAGATAT AGATATAAAA TTTTTTTTTG
26401 AGACAGAGTC TTGCTCTGTT GCCCGGGCTG GACTGAGTGG CATGATCATG
26451 GCTCACTATA GCCTCTACCA TACAGGCTCA AGCAATCCTC TCACCTCAGC
26501 CTCCCAAGTA GCTAGGACTA CATGCATGCA CCACCATGCC TGGCTAATTA
26551 AAAAAAAATT TTTTTTTTCT AGAGATGGGG TCTCTGTGTT GCCTAGGCTG
26601 GCCTCAAACC TCTAGGCTCA AGCAGTCCTC CTGCCTTGGC CTCCCAAGGT
26651 GCTGGGATTA TAGGAATGAG CCACCGTGCC TGGCCAGTAT TTGTATTTTT
26701 GATACTGAGC TTAATTTTGG CAAGTGCTGT GCAAGGCACA AATTGTGGTG
26751 TTAGCTTTCT AACTATTTGG TTGTAATTAT TTGTTAATAT CTGTCTTTCC
26801 TAGTCATGGA AAGTCCATGA GTGCAGGGGC TGTGTCTGTC TTGTTTACAA
26851 CTATATTTGC GCTGGCCAGC ACAGTGCTTT ACACATAGTA GACATTCAGT
26901 GATGTTTCTG AACAAATGAA AGTCCTTCGC TGCAAGAAGA ATTTATTTCA
26951 GTTTAATAAG TTCTTTCTGA ATGCTTCATA TGAAGTAGGC AAAATTTGTA
27001 TTCACATATT AATACAGATT ATAGGTATGG CACATTCCAA TGTCTATTTA
27051 ATTAGTAGAA ATCAAAAAGA GTTTTTATTT GCAATTTTGC CTTGGAGGGT
27101 ATTAGAGCAG ACTGTGCACT AAACAAATAG CTGTACAGAA TTGGCAGCAA
27151 TAATCCCAAA GTTTGAAGGC TGTAGGTTAA TTAGGATAAC TTGACAGGAA

FIGURE 3-8

27201 GTGAGTTAAT CAAATTTGAG AGTTTAATCT TCCAATAATT TATGTTCAGA
27251 CATACTTCAA GTATATCAGC AGGTAACAGG AACTTTAGTT GCAGAATGCC
27301 CCAAAACACA AGAACTCCAG TGGATTTTCT GGCTTCCAGG AATGTTTTGG
27351 AGGAAGAAAA ACCAATAAAA TGATTTGGGG GTCATTTTGT TCCATTACTC
27401 TATATTAAAT ATACTAGAAT TTAAAATATT AAATTTTAAA AGATAAAAAG
27451 ATGCAGTTTA CCTATTAACA AATTAAATAA TTTAGGAATT CTACTTAGTT
27501 CTGTAATACT TTAATATGAG TAAATATGGG CATTTCTGTG TTAGCTAGAA
27551 TTAGATAGAG TATTGCCATT TTTTTCAACT GGCTTATGGT TAAATGGAAG
27601 TAAAGGGGCA AACTACACAT ATAAGAATTA GTAGTACAAT ATTTAATACA
27651 CCCCTGTAAG AGTTATCACA GTGTGTCCTC TGTGAAAAGT AAGGGCTCCA
27701 TGTGTGCTTG TGAAAAAGGC CACTGGAGGC CCTTTTCAAA AATTAAATCT
27751 GCCTCCAGCA AGGTGTTTTT CGATCACATG GAAAGGGGAA GAAGAAGCTA
27801 TCAGGAGCTC TGGGGTTTTT TTTGTTGTTG TTTGTTTGTT TTGCCACTTT
27851 TAACTCTCAA GCTAAAACTG GGGTTTCATT TGAGGAACCA GTAATAGAAA
27901 ATTTCTTATG TACATTCAGC AAAATCTAGT ACTGAGTGGT TACTTTGGCT
27951 TTTCATTGTG GGGATTGTGT GTGTGTGAGT ACATGCACGC ACTTGTGTGT
28001 TTAAGCGTGT AAGGCAGACA GACAGTGGGT ACAGGTCTTT GAAATGGACT
28051 TCTTGGCAAA AGTAATAGAG AAAAAGAGGA ATACAAATAA GGGAGGAGGG
28101 ACAGGGAAGA GCAGAGTCAC AGGAAACAGT GAATGAGGCT GCAGTCTCAG
28151 TCGCCCTTTC TTTGTCCCTC CAGTGTTGTT GCCTGTCTTA TGATGATGCT
28201 GGTTTTCAGC CAACCTTGAG TGAGTAAAAG CCGGGTCTGA GGTCTCAGTG
28251 CCTGCGTGGC TGATATGAGC AGCTTGCATT TCTGACTGGG CCCTGGAGCA
28301 GCAACAGCAC AGATTTCCAG GAACAGTTCC TCTTGTCATT TTTATTCCTG
28351 AGTCATCAAA TTTAGTTATT CAGACGTCTG CTGTTCCCAG CTACATACAG
28401 ATCAAACAAG CAGGGAATTT TTTTCTTTCT CTTCTCTCCC TCTTTTTTTT
28451 GTATTTCCAT CTTGTTTTGT ATACCTTTTC TTTGTTTAAG TCAAGCATTT
28501 GAACATCACT AGTTACCATT TCCTTTAGCA AGCATAGGAC TTCTGTCTTA
28551 CTTAAATGTC TTCTAATGCT GTGATGTGTC ACAGTTAGTT GAGACGTTAA
28601 AGATGTTCCA TACATGTGAC TACATTGGTA AATCTCAAAA ACATCATATC
28651 GAATGAAAAG ACAAAGTTGT GAAATGTCTG ATGTGATGCC ATTTCCTTAA
28701 AAAGTCATGT AAAGCAATCC TACCACATCT CATAGAATCT AAAAGGCTAT
28751 ACTGATGCTA AGATGCACTA TTATTTTCTG TACACTGAGA AAGGGGGAAA
28801 ATTGCCACTT AAATTGTGAT GTAATGTCTC ATTATGGATT GTAAGATACA
28851 TCTACATTTT AGAAATGGTA TAATGTTAAA AATATGCATT TTAAATTGAA
28901 GGTAAATTTA ATTAAATTAT TTCAAAGGAA ATAAGGTAAA TGTATTTTAT
28951 TGAAGCATAG TTATGTAATA AAAATAGAAA AGCATGCATA GGAATGCTAT
29001 TTAGCCAATA CAGGATGTGG TAATCTCTAT AAAGGGAGGG AGGGAAATGG
29051 AGGGGGGAGG CCAGAGGAGG GGCCTCATCT CTGTAGTTTA TTTTTTAAAA
29101 TTATAAAGCA AATATTACCA GAGTTAAGAT TTTACAAAAT TCATTGGTAA
29151 GCACCTATAA TTTTCTGAGT GCTTTCAGTA TTTCATAATG AAAAGTATGT
29201 ATTTTAAAGG TACGTTATCT ATTTATTTTT ATTTATTTTT TTTGAGACAG
29251 TGTCTCATTC CATCGCCCAG GCTGGAGTGC AGTGGTGTGA TCTTGGCTCA
29301 CTGCAACCTC CTTAAAGCTA CATTATTTAA AAGTCACATA CAAAGCAAGT
29351 TGCAGAAGCC TGTATGTAGT GGATTCTATT TTTTTTAAAT AGTATTTAAT
29401 TGTATGTTCT TCTACACTTT TTCCTATGTC CATCTTACCA TAGCTGTGCC
29451 TTTTTTGGTG GAAGTGAGGA CAGATTGCTT TCCACATCTC CATTTTTGTG
29501 TCTGAATTAA AAGATGGACA AGTATCATGT ATTATCTTAG TAGTCATCAA
29551 ACAAGGAAAA AGGTTTCTTT GTTTGCTTGT TTTTTTTAGA TGAAGTCTCC
29601 GCCCAGGCTG GAGCGCAGTG GCACGGTCTT GGCTCACTGC AACCTCTGCC
29651 TCCTGTGTTC AAGCAGTTCT CTGCCTCAGC CTCCTGAGTA GCTGGGATTA
29701 TAGGCGCCTG CCATCACGCC GGCTAATTTT TGTATTTTGA GTAGAGACAA
29751 GGTTTTGCCA AGTTGGCCAG GCTGGTCTTG AACTCCTGAC CTCAGGTGAT
29801 CCACCTGCCT TGGCCTCCCA AATTGTTGGG ATTACAGGCG TGAGCCACTG
29851 TGCCCGGCCT GGAAAAAGTT TTTAATGGTA AAGATGTCAT GGAATGAATA
29901 GGATTGGCTG GCATTATTTC TTGCTGTTAA TAAGCAGTGA GAAATGTTTC
29951 CATTATATGT TTCTTTGAAG CCAGCTTTCT GGTTGCTCCC TTATTCTTTC
30001 TTTCTCAGGC ATGTGGTATC TAGAAAGGGT CAGGAGTACC TTGATAAAAA
30051 TTATTGTACA AGTTGAGCAA AACTCAGTAG TATCATGCCT AGAGATCTGA
30101 TAAAGAGGCA CTTTTAAAAT AGAGCCTTGA AGCTAACAAA CTTTTTTTTT
30151 TTAACCCCTT TACTGAAACC TAAAAAGAGG TCTGCAGTTT TTCCCTCCTG
30201 TCACCAGATG AAAGGGCTGT AGTAGTGTTT GCTTATTCCT CAGGCAGTGA
30251 GGAATCAAAG GACTGAAGGG GTGTTTGTTT TATCATACTA CTTGTAGGGA
30301 CCCCTTACCT CCTATACCCA TGAAAAAGGA ATATTTCTTA TATCCCATAA
30351 TATTCTTTTA GTATCACACT TAGGTTTTAA TGTCCTTACT GTTAGAGTAA
30401 AATAATTTGG GCAAGACCAA TTTTTTAAAT GGCAAATATA GTCACATCAC
30451 TTGATTCAGA ATCTCACTCC CTAGCATCTC GCGTAATGAC TCATAAGAAA
30501 GAAAAAGCTA AATGCATGAA GAGGTTCACT ATACCATTAC TATAAAAAAG
30551 TAAAAATTTG TTTTTGTTAT TTTTTTTTTG TTTTTTGAGA TGTAGTCTCT

FIGURE 3-9

30601 CTCTGTTGCC CAGGCTGGAG TGCAGTGGCG CAATCTCGGC TCACTGAAAC
30651 CTCCGCCACC TGGGTTCAAG CAATTCTCAT GCCTCAGCCT CCGGAGTAGC
30701 TGGGATTACA GGCATGCACC ACCATGCCTG GCTAATTTTT GTATTATTAG
30751 TAGAGACGGG GTTTCACCAT GTTGGCCAGG CTGGTCTCGA ACTCCTGACC
30801 TCAGGTGATC CACCCACCTC AGCCTCCCAA AGTGCTGGGA TTACAGGCAT
30851 GAGACACTGC ACCTGGGCAA AAAAAGTAAA AATTTGAAGT CAACTTAAAT
30901 GCCCCCAAAT ACAGGAATGG TTAAATGCTG ATGACCTACT TTATGTACTA
30951 TGATGCAAAT GCAAATGATA ATGGTTACTG TGCAGCAACA TGGAAGAACT
31001 TGGTGAAAGC AGGATGTACC ACATGACTGT AACTCCAAAA TGTGCTTGCA
31051 TGTGAGGAGA GATCAAGGGA ATATAGAAAA ATGAAGACAT AGTTGTGTTA
31101 GGGTGGTAGA ATTTTGAGTG GATTTCTCCC CCCACTATTG GTAAAAATTT
31151 TTGTACTTAA TTCGTTGTGG GCAGCCAGAT CTTTTAAAGG TAAATTTGAA
31201 TTTCTCTTTA AGAAAATGGC AGACAGAAGG ATGGGGGATA CTAGAAAACT
31251 AAAAGTAGTG CCCCCTTTGA AGATAAAACT AAAACATTTT AAGCCTGGAA
31301 TTGCTTTAGC AGTACATGTA TTGATTATTT AATTTTGTCC TTTAGAAGAA
31351 AGTTGGCCCA ACACAATTAC ATGGAAGTTG GGTTATTGAA GAGGATTGAT
31401 AAAAGAAGTG GGAAGGTCAG GCCAGGTGTG GTGGGTCATG CCTGTAATCC
31451 CAGCGCTTTG GGAGGCCGAG GTGGGTGGAT AACGAGGTCA GGAGATCGAG
31501 ACCATCCTGG CTAACACGTT GAAACCCCGT CTCTACTAGA AATACAAAAA
31551 AAATTAGCCC AGCATGGTGT TGGGTGCCTG TAGTCCCAGC TACTTGGGAG
31601 GCTGAGGCAG GAGAATGGCG TGAACCCGGG AGGCGGAGGT TGCAGTGAGC
31651 CCAGATCAAG CCACTGCACT CCATCCTGGG CGACAGAGCG AGACTCCGTC
31701 TCAAAAAAAA AAAAAAAAAA AAAAAAAGTG GGAGGGTCAA AGCCAATGTG
31751 CACGTTTATT CTTTTGTACC ACCAAATAAA CCGAATTTTT GGACCAACTT
31801 TACTTTAGTC AATATTTGGT TATTTGCTTT GAAGCTATTT GTTTGCAATA
31851 ATGCAATTAA CATCCTCAAC AGAGCCAAAG TCTGACCCTG AAATGGCAGC
31901 ATCTTAACTA ATTTTATCAT TTGAATAGAA ATTGAACCCT TTGAGGGCAT
31951 GAAGTACTAT TTCTTTTCTC TATACTTCTT GTTTTTGGTT TTTGCTTCAC
32001 CACCGATTTA TGACTACCAA AACCCCACAG ATCCTGTAAT TAACTCACAG
32051 CTGCCATCTT GCCACCGTAG ATAGAATTCA GGAGCTGGTG ATGAATGGGG
32101 AGAGAATAGA AAAACTAATA GACTGTAAGA ATTTTAGACC TCTGTGGCCT
32151 TGCTGAACAA TTAATCCAGC CCCTTCACTT TACAGGTAAG TAAACAAGTC
32201 CCAGGGACCC AGCAGTGAAC CTGTGTCTTA GGACTTCAAA TCTAGTGCAC
32251 TTTCTGTTAT ACTTCAAGGA GAGAACTGGA GGGAGAGCGA AGCCAAAATA
32301 CTGACATTTT AGAGGCTGCT TTTTAAGAAA GGATAGGACA TTGGACTTGG
32351 TGTCATTATC ATGTTAAGTT ATAGAATTTT TAAAAACACT ACCACTCCAA
32401 AACAAAAATA GGTAAAAGAT ACGAACAGGC AGTTCACCAA AGAAGAAATA
32451 CAGTAGGCAT GAAAGAATGC ACAACATAAT TTAAGAAATG TAAATAACAA
32501 ATATCTATTG CTATCTCAAG AGAAAACATT TACTATGAGC TTTATTATTC
32551 AAACCATTGG CCTCATAGGG CTTCTTGTTG ACAATCTACC CTAATTTTAG
32601 GAAATGTAAT TATAACATAA AAAAATGAAT GCAAATGATA AAAGACGAAG
32651 TATAACCCGC TCAATAAATT TTCTGAAAGT TATAGTAGTT TTAAACCTCT
32701 TTTTATTCAT TCCCTCCAGT TCTGTCTAGC ACCTGTAAAG ATGAATTATT
32751 GGCTGGGTGT GGTGGCTCAC GCCTGTAATC CCAGAACTTT GGGAGGCCGA
32801 GGCGGGCGGA TCACGAGGTG AAGAGATCAA GACCATCCTG GCCAACATGG
32851 TGAAACCCTG TCTCTACTAA AAAATCCAAA AATTAGCTGG GCGTGCTGGT
32901 GCATGCCTGT AGTCCCAGCT ACTTGGGAGG CTGAGGCAGG AGAATTCCTT
32951 GAAACTGGGA GGTGGAGGTT GCAGTGAGCC AAGATCGCAC CACTGCACCC
33001 CAGCCTGGCA ACAGAGTGAG CCTCCGTCTC AAAAAAAAAA AAAAAAAAAA
33051 AGGTGAATTA TCTCAGGTTT TATATGGCAT TGATGGACCA CTATTGAAGG
33101 AAGCTGTGAA GGACTCAAAG CAGTTGAGAA TAGGTATCGC CACCTAATTT
33151 TACTTGCCTT TCCTGCCACA CACAACCAGA GAAGCTGGTT CATCTCTTAC
33201 TGAGGAAATA TTTTCATGCT TATTCAGAAG ATTTTGCAGG ATCCTCAGGA
33251 AGACAAAGCC AAAGAACAAG AACAGTCAAA GTAGGAACAG GACCAAAAGT
33301 CTGCACAGTC ACTGGAAGGT GTCATGCTGA AGAAGGCAGG GATGGGACTT
33351 TGAAATGAGG CCAAGTGCAT TTCAGTAACT GAGTGGGTTA TCTGTTGTTG
33401 GCAAGCAATG TGCCATACCT TTGAAGGGCT AAGCTAGCCC AGGAGTTCTG
33451 ATAAGAGCTT TGAACTTAAT ATCGCTGACA CAAAAATAGG TCTCCTCAGA
33501 TCCTATTTGG AGGTAAACCG GGTAAAAGTG ATAAATAGTA TTGCTTTTTA
33551 AAGTTCATTA CATATTTTAG AGGTGAGTCA GTCAGACCAT ATAGAGATGC
33601 TTCATTTTAA TCCTCTGTAG CAAGAATTGG CAAAAATATT CTCTGTAAAG
33651 AGCCATTTAG TAAATATTTT AGAGTAGGCT GCATAGTCTC TTTTGCAGCT
33701 ACTCAGCTTC ATACAACAAG TAAACAAATG AGTGTGACTG TATTTTAGTA
33751 AAACTTTGTT AGACGCTAAC ATTGGAATCT TACACAGTTT TTATATGTTT
33801 TGACATATTC TTTTGATTTT TTCCAACCAT TTAAAAATGT AAAAAGCATT
33851 TGTAGCTTGT GAGGTGTACA AAAATGAGCA GTGGACCATG TTTTGCCTGC
33901 AGGGCACATT GTGCTGATCT TTACTCTATA GCATCACTGC CAGTAGAAAT
33951 ATAATGCTAG CCATATGTTA TGGACTGAAT GTTTGTAACC CCCAAAATCC

FIGURE 3-10

34001 ATATTTTGAA GCTCAAACCC CTGGTGTGGC AATATTTGGA GATGGGACCT
34051 TGACAGATGT ACTTAAGGTT AAATGAGGTC ACAAGGGTGG GGTCCTGGTC
34101 AGATAGGATT AGTGTCCTTA CAAGAAGAAA CACCAGATAC ACTCATTGAG
34151 AAAATGCAGA ACATTGTAAC CTTACTGGTT TAGAAACCAT AATACTGGCA
34201 GGCTTGCTGG TCTTCATTTT TATTATTTTG AATGTCTTCT TTCAAGTATT
34251 TTATGTATCC AGACCCAATT TGTATGATAA GGGAATTTCT TTCAAACTAT
34301 TAGGTGAAGT CTTTAATTAT GAAGTTGGCT TAAGCCATTA AGTAGTATTA
34351 ATGGGGACAT CCATCTTAGA AATTAACTGG AAATGTAGAC TAGAAATATA
34401 AAAAAGTGTG GTCATGACTA ATGTCTGTAT CCATTTCCCA AAAAAGAGAT
34451 TTGCGTATGT CCTCACATTG CAGAACCACA CTGACACTCA GGGAAACATG
34501 CTGACGTCAT TCTTCCAACA TCCTTAGTTT GGATTTCATG AAACATTTTT
34551 TTCCATTTCT CCTTTTTCTG TGTTTAGTAA TCCTTTCTTG GATTCTTGAA
34601 ATCAGAGGCA CTTATTGGAG TTGATAAATG GCAGTTCTTC ATGTCACTGT
34651 TTAGAAATTT AACTTAGCTT ATGGACTAGT TCAGGCTTAG ATAACTCTGG
34701 AAGTCTTTGC ACTTTCAAAT ATACTTTTTT ATAGTTGAGT TTCTACACAA
34751 TATAACATTT TATAACTTAA ATGGAAGTAA AGTGAAATGC CAAAATGCCA
34801 GTTGATTCTG TTAATGTGTG CATCCCTCCA TTCCAGAATT CCAAAAATAT
34851 CAGACATAAT GACTTTTATT TGTATTTTAT ATTATTAGAT TTATCAGATA
34901 TAAAATTGTT CTTTTAAACT GCTATTTCTA AACACTCTTT GTTTCTTTAG
34951 TTATCTTGTG AACTGGTAAT AGTTTTCAGA CTGGTGCCTG TCCATGAACC
35001 ACTTTCAGCA TTGGTCCAGA TGACTTTCAC TTATACCATG AGCCAGCAGA
35051 TACTGTAGAT TATTTGGCTG TTTTTAAAGA GAACATCATA TCATAGTGTT
35101 GGTTATTAAT ATTTGAGCAA GATATTAGGG TTGTCTCGAT GTCCCAATCA
35151 TTGTGGAGTG CTTAGTGTGT GTTTGTTTTT TAAGAATTGA CTTGGAAACA
35201 TTTATCTTAA GGTGTTAGGG TTATTTTTAT TCTTTAGTAC TAACGGTATA
35251 AACTGAATTC CACTGAAATA GAGCTTATTT TGATGTATAA TTTTTGAAGT
35301 TATTTTTTAT AAGTCTGTGG GTAGATAGCA GCCAAGTAGG CACTTTATTT
35351 CCCTAATAGA AAAAAATATA TATTTTTGTC AAATATTTCT GTTTTATTCA
35401 TTCATTCAAA GTATATTTGG AATGTTTATT TCCAGGAAAT TTTGGAATAT
35451 ACAATACAAC CAGCTTCTTA TAACTCCACT TTAAGTGAGC CATAGGTCAA
35501 ATAATGACCA GCAAAATGTA ATGACACGTG TGCCTCTTAC TCCCTGTTGG
35551 AGGAATTGAG GCACTCTGGT AACCCTGTAG GCCTGGATTA GTCCAGTTCA
35601 TTGGCAGCAG CATTATCCAG ATTTTATTGT GGCCGGCAAC GGTGGCTCAC
35651 ACCGGTAATC CCGGCACTTT GGGGGGCTGA GTTGGGCCTG TTGCTTGAGC
35701 CCAGGAGTTC AAGACCAGCC TGGGCAACAT AGGAAACCCT GTCTCTACAA
35751 AATATATAAA AATTAGCTGG GCGTGGTGGC GTGTGCCTGT AGTACCAGCT
35801 ACTTCGGAGG CTGAGGCAGG AGGATCACCT GAGCCCAGAA AGTTGAGGCT
35851 GTGGTCAGCT CTGATTATGC CACTGCACCC CAGCTTGGGT GATACAGTGA
35901 GACCCTGTCT TAAACAAACA AAAGAGATTG TATTGTGTTT TGAAAAACAT
35951 AGTTTTTAGA GAAATTCTGT GTTATTCTTC TATTTCCATT ACAGCTTGGT
36001 TTAAAGAAAC TATGGAAAGC TACAGACTAA CCTCCTTTTC CTTTGGTTTT
36051 TTTTCCATGT TATAGCAGAG AACCAGCCAG GGCCAGATTC CTGGGCTTGT
36101 GTCCTCTGCA GTTGCACAGG GCTCTGCTCA GAAGGGTCCC ATACTTGGCT
36151 TAATGCTCTG CTGTCAGCAT CTTGAAAATT TTTAATAATT TTATATTTCA
36201 TCTTGTTTTG TAAGTTATAA GCCCAGTGGG CCAGTGGAGG CTGTTCTAGG
36251 GGTCTGGGTT TGGCTCATGA GAGGTACATC CTTCCCACCT CCCCAGGATG
36301 GGCTCTCAGC AGCTGGCTTT CTGTACCCTG GCACCCCAGG CCCTTCTTGG
36351 CCCCTACCCC GTGTTCACTG TTGCCCTCAC CTTGGGTGAT GACTGGGTCA
36401 CATGGGAGAG GAGGAAACCC ATGTTCTGCT GTCACCCTCC ACCTCTTGCT
36451 GGGTCCTGGG TACAAGCTCA GGGAGGGTTG GGATCAGGCG TCTGTGACTC
36501 ACAGTGTTGT GGGGAGTGAT GGTGGTCATT TTATCCCAGG CTGGCAGCGT
36551 GCTGGCATAT TGATGAGCAG CTGCTGGCAT GATGATTTGT CCTGACCCTG
36601 GACCGGACTC TGTATCTTGG CTTTAGTCTA GTCCACTTCT CAGCATTCCT
36651 TGGAAGCCCA CCTAAACTGC AGAGACAAAG AAGAATATGC TTTGGGATTT
36701 TAAAGATGGA ACCCCATATG CTATCAATTG GAAAGGAGTT CAGATTCTGG
36751 CTGGGCTGAA TCACTGATTT TCCAGGCAAG TGGAACTTTA ATGACCCACT
36801 AGTGCCCTTT GTTCACTGTG GGCTGGCTGG GGAGGAACCC TTGCCTTTAA
36851 AGAACCCTGT GCTCTTTAAA GTTTCCAGGG CACAATTGAG ATCCTTGAGG
36901 TACAGTTCAT CCATGCAGAC CTAGCACCTG CTTTAGGTAA GGGCTCTTGT
36951 TACTTCTTGT TCTTATCGGG TTCTTATGTT CAGCTCAGCT TCCTGCTTCC
37001 TCAGAGAGAA GTGGTAGGTT CTGTTTTAGG CTTTTCTTTT TTGGTGGTGG
37051 TGGGAGGTGT CAAACCGCAT TATTGAGGGG AGTGGGGGTG GTAAAGCAGT
37101 GGTGAGGACG ATGAGAAGCG ACCTCCTTAG CCAGGCCAGG AGCTCACCTC
37151 CCAGACCCTG CCCTTCAGGC ACTGAAAACT GTTTTAGGCT TTTCTGCATG
37201 GTTGGAGTTT ACTCTTAGTT GCTGACAGTC ACTCATCCCC TGGAGAGGGG
37251 TACTGAGCAT AATCACTGTT GGTATTTTAT TTACTACCAG TTCTTCATCA
37301 TTACCGAGAC TCATTCCATT TGGAAAATAA ACATTGCCTC ATCTGTTCAT
37351 GGTATATTTT TACACTTGCC TTTCACTCAC ATGTTGCATA TTAACTGAAA

FIGURE 3-11

37401 AGATCAGCAT GATCTTGGAG CCTAATGTGA TCTGGATCTT ATTACCATGT
37451 ATGTTTTTAT CGTCAGGTGC TGCTTTGGAC TCTGGCCCAT TCCTCCCACC
37501 TGTCTTCATC CTGTCGGGGT TGATCTCTGC CTGGTTCCAT AGTCACTTGG
37551 CTTTTGGAGT TTTATTTCTT CTCTAACCCA CTCTTCCCTT CCCTGTCAAA
37601 TATAATTCTC TCATGAGCAA CTGTGCCTGA TAATCCCAGT GTATAGGCCA
37651 TAGCTCAGGT TGTTTTAATG CTTGCAAGGA AGTCTCTTCA GATTACTACA
37701 GGGTAAAAGA ATCCATTCCC TTTCAACCTC TCATTACTTA CTGAATCAAT
37751 TATTTATTTA CTTGTGTAAT TCACATCCAT GCTCAGGTAT GTGTTTAGTT
37801 TACCTCTGAG ACTGTGGCTC CCTCTAGACA AGGGTCTTTG TCTGGTATTA
37851 GTCCACTTCC TCAGTGTCTA ATTTCATGAG GACAGAATCA GGGCAAGACC
37901 TACTGGCTAA GACTTGCTGC TGCCTCAGTT ATTTGGTTTA CTGTTTAAGC
37951 CCTTGAGCCA GGAGGTGACA GTATCTTTTG TGTTCCAAGT CAGATAACCA
38001 GTAACAGTAA TATTCTGATG TAGGTCTAAG GGGCAATAGG AGCCTGAATC
38051 TGAGCCCCTT GGCAGGGATG GTTTCCCTGG GTTACTGGAT TTGAGTCTTT
38101 GTCTTCCGTA ATAGTAACTT CTGTGACCCT TGGCTCAAGG AGTCCTTTCT
38151 AACCTAAATG CCCTCCTGAA GAAAGTGCTG ATATTCATCA GAAAAAAAAG
38201 TAGGGTTTGT GTTATGTTAC CATCTGGGAG ACATTGGCTT TAATCCTCTC
38251 CCCTCTTCTT TTCCTTTTAA AATTTTCAAG AATCCCCAGA ATTGAGGGTC
38301 AATCCTGGAT CAGAAGGGGC AGAAAAGTCA GCCAGTCAGA GGTGAGAAAG
38351 GGGTCTTGTG TTCCAGTTGC TCTACTTTAG AGGGCTTGAG GGCCGCATTA
38401 GGAGCAAACC AAATACTCTG AGGTCCTAGG AGAAGATTGC TCTAAATCCA
38451 GGATGATTTC ATCATATGAC TTGAGATGGA GGGGTTCTTG ATTAGTAATG
38501 GGCTTTTTGA TTGGCAGGAA GTAGGTAAAC TGCAGGCAAT AGGATTGGCC
38551 ATAGAAATAT AGGGAACTCT ATGCTTTCAA CTCTTGTGAA GTTAGGACCA
38601 AACCAAATTA TTTGGGGTTC ATTGGTGGCT AAGAAATTAA CTTTTATACT
38651 TATAGTTAAA AAAAAAAAAC TTAAAGAACT TTTATTAATT CTTCCCCAAT
38701 AGAGTGATTT TTTTTTTTTT TTAAGAGACG GAGTCTGGCT CTGTTGCTCA
38751 GGCTGGAGTG CAGTGGTGCA ATCATGGCCA CGGCTCACTG CAGCCTTGAC
38801 CTCCCAGGCT CAAGTGACCC TCCCACCTCA GCCTCCTGAA GTAGCTAGGA
38851 CTACAGGCGT GCACCACCAC ACCCAGCTAA TTTTTGTATT TTTTGCAGAG
38901 ATGAGATTTT GTCATGTTGC CCAGGATAGT CTCCAACTCC TGGGCTCGGG
38951 CCATCTGCCC GCCTTGGCCT TCCAAAGTGC TGGGATTACA GGCATGAGAC
39001 ACCATGCCCA GCCAGAATGA TTTTCCTTTG TTGAATTCTC AAAATTTTAT
39051 TCTGGCTTAG GTAATTGAGT AGGAATGGCA GAGAGTAATA ATTTTCAGTG
39101 TAATCTTTAG CATTACATAA ACCTCTACAA ACATGAAGTA AAAAGTTCTT
39151 TACAGACTTT CTATAGGAAA AATAAATGTT TATAGATCTG CTTATCAACT
39201 GTTTGTTCTT CCTACCTCTG CATTTTCATC TATTGATAAC TGTCAGAGTT
39251 CAAGGTCAGT TAATTGTACA TTTTCTGGGA GTTTTTCTCA CTGTTAATTA
39301 GTAAGACTTT GAATCATTAA GTGCTGATGA CTACATGGCA CAATTCATTA
39351 AAGACAGTCA TACTTTAACC TTAGTACAGC TATAATGGGA CGGTGAAATA
39401 CAACTAGGAA AGTATCTTAT TAAAGTATTT TACATTTTAC TAATAAAATT
39451 TATTTAAACC CTGAATTAGT GTTTATTTTT GTTTCCGAAT TCTCAGTAAC
39501 ATCTCAGTAG ATCCGTGAAT CCTGCCAAAA GACGTATTTT AATCCAAAGA
39551 TCTTTCTGCA TCTATAATTT ACCACACTAA AGCTCACATT ATCATTAAAA
39601 GCAGCATTAT CATTAAATTG TAGACATTTA TTAGTTCATC TCCAGGCCTA
39651 ATGGGAGTGG TTTGGAGTAA ACCTTTGTAG AAACAATAAT ATGTTTTGAT
39701 AAATGCATAT GGGCGTGGAG TGGTTATCAC TTACCTCCTC ATGAAATTTC
39751 TGGAAAGGTG ATCCGGAAAA CCAGGACACA TTTATGGTAA GATATAACAC
39801 CTTACCATAA CGATTAAGGC TCATTGAAAA ACTCTGACTT AATGAGTTTA
39851 CCAAAAAACT GTTCTATCCA CATCTCATCA AACCGCCCCT GAAATTCCCT
39901 TGCCTCTGTT AAATTTTTCT AGACAGTCTG AATAGAGGCA TGTAATTTTT
39951 TTGGATTTTT CTGTGGTTAA ATAAATATCC TTTACAACTC TCTTTATTCT
40001 TGAATATCCA TAAGAGTTTA TATTTATACT GTATGTTTGT TATTAGGATT
40051 CCTTTCAATT GCTATATAAA AAATGTAAAG TCTGTTTACT GCCTTAAACC
40101 TTCTGGTGTA TTTTTATATA AAGTAACACC CTTAATTCTA ACTTGGCCAA
40151 CAGGTAGGAT GGTATTATTA TTATCTTCAT TGTACAGATA AGGAAACTGA
40201 GGCTCAGATT GACTAGATCA AACAGGAGTT TTCTGGAAAA CCTAGGACAC
40251 AAGCCTAAAT CTTTGAACTC AAATACTGCT CTACACTGAA TTACAGTTAT
40301 ATACTGATTT CTGTTGTAAA TTCTTAGAGA AGACAGACAT AGAAATTAGT
40351 AACTTGAGTC AGTAGCGGCT TTGTTCAAAC ACAGGCACAT GCATATTTTA
40401 TGGTATATGT TTATATCTGT GTAATACTCA TCATAAATGT CAGATTTATA
40451 ATCGATAGTG CCCATTTCTA AATTTATAGT TGAAACCTCT GATAGGAATT
40501 AGGAATTATC TTAAAGTCCT AGAATAATAA ATATTAAGAT TTTGAAGACT
40551 GCTTAAAACA GTGTTCCAGG CTGGATTTTT TCCTTAGTTT TTTTTTCCTT
40601 AGTTTTTTAC CTTAGTTTTT TCCATGTAAT GAATGAGGAG CCATTGGGGT
40651 TGGTGTTTGG TTGTAAAGGG CAGGTTGACT TCACCAGGAG GTGTTTCTTG
40701 GTATTTATGG ATCCTCTTTT TCACACTAAT CCTTTGATTA GCTTCTTCTT
40751 ATAGTTCATG CTTGTACAGT TGCAGCTGAA TGGTAAGTAG GTAGAAATAT

FIGURE 3-12

40801 GCATTGACTA ATGTTGAACT ATATCTAGGA GCGTCATATT CATGCTACTA
40851 CTGAGCACTG TGACCTAGGT GAAGTGTGAG ATTAAAGAGC TTTATTGCTT
40901 ACCACGTTTC TTTTTGATCT AGTTAGAGTC ATGTGAGGCA GTCCATGACT
40951 TTTAGGATGT TATAATATAA TCCACATAGT ACTTTATACT TTTCAATTTG
41001 TTTTTATATA TTTTTCCTTA CTATAAGCCT GTCAAGTCTA GTGAGATGAC
41051 ATGGTTATTA ATATATTACA AATGAAGAAC TCTGACTTGG AGGTAGTATA
41101 GCACTGTGGT TAATAATAGA AGTTCTTATA TTTTTATATG AATTATATAT
41151 TATGAACCAG ACCAGAATAC TAGCTCTACC AATTACCTAG CTGAGTGATG
41201 GTGCAGCAAC TTAATCTCTC AAAGCCCTGG TTTTCTCATT GTAAAATGGA
41251 GGACCTTATA GGATTGTTGG GTAGCTTAAA TGAGATAATA GGTAGAATTA
41301 AAAGTAACGG CAAAAACCGC AATTACTTTT GCACCAACCT GATATTTAGA
41351 ACAGTGGAGG CATATAGTAG GCATTCAATA AATATTGGTA GTATTATTTT
41401 CAGAGGTTAT TACTCATTTA ACTCAACAAA AATTGAGAGC CCCTTTCTAT
41451 ACGCTAGTCA TTGTGCTAGG CATCAGGAGC ACAGGGCAAA CCTGGTAGGG
41501 TGCTTACTCT GATGGAATTT ACATTTTAGT GGTAGAAATG GTAAATGAGT
41551 CAACAAGAAT TTTATGGAGT GACAAGGACT TTGAGGAAAG TAAAACGGGA
41601 CAATGCTTGA GAAGGGTGCT AGTTGGGAAG TCCTATCACA CGAAGTGGTA
41651 TTTGATCTGA GACCTGAGTG ACCATAGCCA CATGAATACC TGGGGGAATC
41701 ACTTCCCAGG AGGAAACAGA AGCACAGAAG CCGAAAGGCA GGATCACAGG
41751 CGGCACATTT TATTGGTGCA GTGTGGAGGT AGAGGGACTG GATTGTGTAG
41801 GGTTTGTAGG CCAGGGTAAA GAGCTTGGAT TTCTTCTGAG TGTAGGATTT
41851 TGAGCAGGGG AGTGATGTGT CTTATTTTGG GCGGGCGCTA GGCCAGGTAC
41901 TGTGAAGGAA ATAATGGTAT AGAAGAAATA TGTCTTTCTT GCTAGAATCT
41951 TATGTGACAT ACAACTAGTG GTGGCTGATC AGCAATTAGA ACTGCAGTGT
42001 TCTTGCTGGT GTACTCCTAA TCTGATCATT GGATGATATT TATCACGTAT
42051 TTTTGCAGCA GTCATCTCTA AAGTGGCATT ATGGTTCTAG TTATCTATCA
42101 TTGGATCTGT TTGCTCTAAC AGTCCGAAGG GGCATGAAGG GTGGAAACAG
42151 AGTAGAGGGA AATGGGTGAA AAGGCTATTT CTAAACACTA GACACAAAAT
42201 AGGGATTAGG AAGTAGTGGG GGTTAGGAGA AAAAGCAAAA TCCTGTTGTG
42251 CATGGAGAGG GACATTGTTG AAACTGTGGC TTCTTGAGAT TACATGAGTG
42301 AACCTGCGTG GACTGGGTCG TCTTCCCTAT TCACCAAGCT GAATCCAGGC
42351 CCAATCCAAG ATTACTGCTG CCTTTGCTCT GAGGCTTTAG ACTGTTAGTA
42401 GCTTTTCTCG TCTCTACTGA GGCCAAAAGG GGCAGTGATA CCCTTGAATT
42451 TTCTTCTTAA AACAGGGTTT CATTTCCTTG GAAGTTTGTT TCTTTGAATC
42501 TTTCTGTCAG TTTAACTGTT ATCATCAATT GGTTAGCATT CTAATAATAA
42551 TTATAATTAT AGTAAACATT TATTGAGTGC TTACGAAGAG CCAGTTCCAA
42601 GCTTTTTTAT CTCCATTATT CTGCTACTTT CCTTCTCATT TTACAGATGA
42651 GGAAAATGAG GCACAGAGTG GTTAATTAAT CTGTTTGAGG TCCCGTAGCA
42701 GGTCAGTGAT GCCAGGGTTC AAACCTACAC TTAACTCTAC ACTAGAGACT
42751 GTTTTCTTAA TTATTTCTTC ACAATCATAT GTTTAATGAT TACTTATTGA
42801 TTATTTAGTG GTCTGATAAG AAGAGGGAGC GGTGCTCTTC TGTTGGAGAA
42851 GAAAGGCTGG CTGATCAAGA CACACTGGTT GGTTTGAAGA AAAAATATAG
42901 ATGTTAATTC CATAACACCA CACTCTAAAC ATTTCTACTG GACGAGTTCC
42951 ACCTGTGTGC CACTCGAAGT CGGATGCAGT AAGGAAGGCT TTTTATTGAG
43001 GAGAGAACGA ATACCTCTGT ATTCAAAAGA GAGTGTGTTG TTCCTTATAG
43051 AAGATGGAAG GGGGCTTGCC AGTGACAGAT TATGATGATT ACCTCCTTAG
43101 TGGTTTTTTT TTATTGCACA GACTATAATA ATAATTATAA AAATTTGTAA
43151 ATCAAGAAAA ATCTTTACCC CAACAAATTG TTTTTATATT TTTCATATTC
43201 ATATTTTTTC ATATTCATTT TATTTTTTCA TATTCAAAAA TAAAACAATT
43251 TTCATATTCA TTTCAATTGT TTATTTTTTC ATATTCAAAA ATAAAACAAT
43301 TTTCATATTC ATTTCAATTG TTTTATTTTT TCATATTTTT CATATTTCAT
43351 ATTTTTTCAT TTTTATACAC AGTTATAGAG TATCTACTTA TTTTTCATTT
43401 AACATTATGC TATAACAGGT TTTAAAGTTA GGCAGTAAGT CTTTCCAGTG
43451 AAAAGGAAAA GACTGAAACA TGAAAAGGTT ACGAGGCATT ACTAATTGAT
43501 TGGAATGATG CCTACCAGGC AGCCCATGTC AGTTTTGGAG CTGGCCTTGG
43551 AAGGAACAGC TGTGTATTTG GACCTTGGAG CAGTCTGTCA GGCTTCCAAG
43601 AAAGTGGGCA AAGAAGGAAG GGGGGTTGGG GGCACAAGTG GAAAAGTACA
43651 GTGGTTTGGA ACTGATGTGA AACTATCAAG CTGGTTGAGT TCAGTGTTAG
43701 AAAATAGGAA AAACAAAACA ATTTCTACCT GATATGGTAC TTTAATTGAG
43751 GTTAGCACTT CAGTGAATAG GCAATGCATT ATAATTTGTC AGAATTATCA
43801 ATATTTTGTT CTTGTCAGTG GCTTTATATT TAATAGTTAC ATGTTATAAA
43851 GCCAGTTTTA TCAGTGAAAG AACAAAGTTC TTGACAAATG CTATTTTAGT
43901 GATAAAGCTG TTATTCTGTA TTTAAATTCA GTTTAACTCA AAGTGGTTTT
43951 TAAGTTTTAC ATTTGTATAA CTGTAGACTA GCCATATGGC ATTCAAAGGC
44001 CTCCAAGATA TAACTTGAGA CACTCTTAGG ATGGAATCCC ATATTCAGAA
44051 TCTTTACTGA CAGCTCCCCA TATCCAAATT GCCTTTTCTC TCTTGGGCCT
44101 CTGGGTGTGA CAAGGGACAA CTTGAGTGAA AGGCCACAGG ACAAAGTAGA
44151 GATGCTGTTA AATCTCAATG CAGAAGGCCT GAGACAACTG ATAGAGACAA

FIGURE 3-13

44201 TTGTTTCTCT TTTCTCCCTC ATAAAATAAG TAAAATTCAA AGTGATTTTT
44251 TTTTAATTTT TTGCTTTGAC ACATTGTGTT TAGTCTGATT GAGGTCATCT
44301 TACACAACAG AATCCAAAAT CCTGAGAAAA TAATTTATAC AACTAGAATT
44351 CCAAAATTAG GTTTTGCATT AAGTTCAAGC TATCATTTCT TTAAGAAAAC
44401 CTACACTGGA GTTACCAACT GGAATATTTG GTTATGAGAC ACAGTCATCA
44451 TGTCAGAGTA TGTAGTTTGC TTCTAACACT TATTTGTGTG CTTTTAAGAC
44501 CTGCTCCCTG CCACACTTTA TCATAACAAA AAACAAGAGA CACATGTTAG
44551 TACTGTAGCA TGTATAGTCA TGGGGAATGT GTTCTAAACG AAATCCCATC
44601 ACAGTGGAGT CAACACAATA TTTCTAGGGG AGAAGGTCTT CCTGTTAAAC
44651 TCATGTGAAG GCATTCTTCT CTACTATGGG AACTTGTTTA TGTGCCCTCT
44701 AGAAGACAGC TGAGATGGTC TTCAGTGAAT CTGTTCACTG ACATGTTGCT
44751 AGTTTCTTTA TGTTTACATG TAAACAGTGC TTAGAAGCAT CTTTCCCCAA
44801 GCTGTTCTTT TTTTGATGGA CTCCCCCTTT TTGGGGGAGT ATCAGTTGAA
44851 TGAGCACTGT TGTACTCTCA GCACTAGACA TTCTGTTATG TGTGTGTGGA
44901 GCTTCCTTGT GGCTTGGTAG CATTACCTCT GCATTACATG CTGCAGCCTC
44951 ACCAGGCAGA AGGGCCATTC ACTTCTGCTG CTACTTACCA GCACCTCTGT
45001 TTCCTCGAGT GTGTTACCCT CCCACCACCT GGGCCTTGGA CCTGGGTCCG
45051 TTGGGCTATT ATCCATGCTC CCACCTGCCC ATCTCTGCTC TGATATAATC
45101 TACTAGGATA ATAAAGTAGC TTTTCCCTAC CATAATGTTT TTTAAATAGA
45151 CATACAGCAA TGTATAATGG TTAAGAGCAT GATCCAAACT GTGTCTGATT
45201 TTGAATCCTG GCTGAGTTAC TAGCTATGTA ATCTTGAGAA AACTACTCAA
45251 TCTCTCTGTG CCTTAATTTT CTTAGGTATA AAGCAGATAC TAATTATGCC
45301 ATCATAGGGT AGGTATGAGG ATTAAATGAG TGAGTATTTG TAAAACACTT
45351 AAAACAGTGA CTGACTTGGG CTACCCCTTT TGGGTCCCCT CCCTTTGTAT
45401 GGGAGCTCTG TTTTCACTCT ATTAAATCTT GCAACTTCAC ACTCTTCCAG
45451 TCTGTGTTTG TTATGGCTCA AGCTGAGCTT TCGCTCGCTG TCCACCACTG
45501 CTGTTTGCTG CCATCGCAGA CCCGCCGCTG ACTTCCACCC CTCTGGATCC
45551 GGCAGGGTGT CCACTGCACC TCTGGTCCAG CGAGGTGGTG CCCATTGCCG
45601 CTCCCAATCG GGCTAGGGGC TTGCCATTGT TCCTGCACGG GCTAAGTGCC
45651 CTGGGTTCAT GCTAATTGAG CTGAATAGAG CTGTAACACT CACTGTATGG
45701 CCCAAGGTTC CATTCCTTGG AATCTGTGAG GCCAAGAACC CCAGGTCAGA
45751 GAAGAAGAGG CTTGCCGCCA TCTTGGAAGC AGCCCGCCAC CATCTTGGGA
45801 GCTCTAAGAA CAAGGACCCC CCGGTAACAT TTTGGCCACC ATGAAGGGAC
45851 TTCCAAAGCG GTGAGTAATA TGGGACCACT TTTGCTTGCT GTTCTGCCCT
45901 ATTCTTCATT GGAATTGGAG GAAAATACCG GGCACCTGTC AGCTGGTTAA
45951 AAACAATTAG CATGGCCACC AGACTTAAGA CTCAGGTGTG AGGCTGTCTG
46001 GGGAAGGGCT GTCTAACAGC CCCCAACCCT TCTGGGTTGG GAGCGTTGGG
46051 CTGTCTGGAA CCAGCTTCCA CTTTCAGTTT TCCTGGGGAA GCTGAGGGCC
46101 GACTAGAGGC AGAAAGCTGT TGTCCTGAAC TCCTGGTGTT AGCTGGTTGA
46151 GATCATGGCG CAGCCAGAAG TCTCTACTCA ACAGTCAGCC ATGCGTGCAC
46201 CCCTACCTTT CCTTCTGACC TATACCTCTT GGTCCTGACC ACAACTTGCT
46251 TGAAAGTGTA GCCCCAAAAT TCTCCTTACC TCTGAATCTA CTTCCTCCAA
46301 TCCCTGCCTC CTAGGTACTA ATGGTTCAGA CCTTCATTTC CTCTAGCAAG
46351 CTGTATCTCC AAAGGGATCT AAGGAAGCTC TATGCTGTGT CCTTAAGCCC
46401 CTAGGCTCTG AACCCAGATA GTCTTGTCCC TGGTGTCCCT CCCAATTTAG
46451 GCATACAGCT CTCAACATGG GCAGTTATGT AGGACCTGTT CCCCACCATC
46501 CTTGCCAGGG TCCCAGGTTT GTAAAGGGCT AGGAGAAGAG AGAGAGAGAG
46551 AGAGACAGAG ACAGAGGGGA GAGAAAGAGA GAGAGACGAA GAGGGAGTCA
46601 AAGAGAAAAA GAAAGAGAAA GATAGAAGTA GTAAAGAAAA AATAGTGTGC
46651 CCTATTCCTT TAAAAGCCAG AGTAAATTTA AAACCTATAA TTGATAATTG
46701 AAGGTCTTCT CCTGACCCTG TAACACTCCA ATACCACTTT ATTGTCAGTG
46751 TAAACAAGGG GGTAGCCCAA AAACACTGAG ACCACTGACA ACCTATCAAA
46801 ATCCTATCAA AAATCCGTAA CCCAGTAACA CGTGGATGGG CCAAAGGCAT
46851 TCAGTCGGTA GCGGCAACTG CTTTGCTAAA AGTAGAAAAA TAACTTTAGA
46901 GGAAACCTCA TTGTGAGTAC ACCTCACCAG TTCAAAACTA TCCTAAGTCA
46951 AAAAAGCAAA AAGGTAACTT ACTAACTCGA AAATCTTAAA ATATGGGGCT
47001 CTTCTGTTAG AATAAAGGTA ACTTATTAAC CACTGAAAAT TCCCTTAACC
47051 CAGCAGATTT CCTAACAGGG GATTTAAATC TTAATTACCA TACAAAGGTC
47101 CGACAGATCT AGGAGGAACT CCCTTCAGGA CAGGATGATA GATGGTTCCT
47151 CCCAGGTAAT TGAGGGAAAA AACCACAATG GGTATTTAGT AATTGATAGG
47201 GAAACTCTTG TAGAAGCAGA GTTAGGAAAA TTGCCTAATA ATTGGTCTGC
47251 TCAAACATGC GAGCTATTTG CACTCAGCCA AGCCTTAAAG TGTTTACAGA
47301 ATCAAAAAAA CTCAATCTCA ATCCTGACTC AAAAGGTTAC CTACACCCTC
47351 TCTGAAATGA ATTTGCATAA GAACTGTTGT TTATGGGAAT GCATCTTGAT
47401 GGGGCAACTG GGTTGTTATG AAATACTCAG GAACCCAGCC CAGCTCCAGG
47451 ACTCACCCCT GAGCAAAAAA GCAATGTTGG GCACTCTGGT AAAGGACCAC
47501 TAGAATCCAG CAGCCTGGAC CCCTTTCTTT GTGGTCAAGA AAGGCAGGAA
47551 AAGGGGTGCA GGACTGCTAC ATCAGTGAGT GCAACTAATC TGATAAGCAG

FIGURE 3-14

47601 AGGTCCATGG GTGGTTACAC ACCCTGGAAA GGAATAAACA TTAGGACCAT
47651 AGAGAACGCT CTAGGACTAA TGCTCATTGG AAAATGACTA GGGGTGCTGG
47701 CATCCCTATG TCCTTTTTTC AGATAGGAAA TGTTCCCCCC AAGGCAAAAA
47751 TGCCCGTAAG ATATATTCTG GAGAATTGGG ACCAATTTGA CTCTCAGATG
47801 CTAAGAAAGA AATGACTTAC ATTCCTCTGC AGTACCACCT GGCCATGATG
47851 TCCTCTTCAA GGGGGAGAAA CCTGGCCTCC TGAGGGAAGT ATAAATTATA
47901 ACACCATCTT ACAACTAGAC CTCTTTTGTA GAAAAGAAGG CAAATGGAGT
47951 GAAGTGCCAT ATGTACAAAC TTTCTTTTCA TTAAGAGACA ACTTGCAATT
48001 ATGTAAAAAG TATGATTTAT GCCCTACAGG AAGCCCTCAG AGTCTACCTC
48051 CCTAACCTGA TGTCCCCCTG ACTCCTTCCC CAACTAATAA GGACCCCCCT
48101 TTCAACCCAA ACAGTCCAAA AGGACATAGA CAAAGGAGTA AACAATGAAC
48151 CAAAGAGTGC CAATATTCCC TGGTTATGCA CCCTCCAAGC GGTGGGAGAA
48201 GAATTCGGCC CAGCCAGAGT GCATGTACCT TTTTCTCTCT CACACTTGAA
48251 GCAAATTAAA ATAGACCTAG GTAAATTCTC AGATAACCCT GATGGCTATA
48301 TTGATGTTTT ACAAGGATTA GGACAATCCT TTGATCTGAC ATGGAGAGAT
48351 ATAATATTAC TGCTAAATCA GACGCTAACC TCAAATGAGA GAAGTGCTGC
48401 CATAACTGGA GCCCGAGAGT TTGGCAATCT CTGGTATCTC AGTCAGGTCA
48451 ATGATAGGAT GACAACGGAG GAAAGAGAAC GATTCCCCAC AGGGCAGCAG
48501 GCAGTTCCCA GTGTAGCTCC TCATTGGGAC ACAGAATCAG AAGATGGAGA
48551 TTGGTGCCGC AGACATTTGC TAACTTGCGT GCTAGAAGGA CTAAGGAAAA
48601 CTAGGAAGAA GCCTATGAAT TATTCAATGA TGTCCACTAT AACACAGGGA
48651 AAGAAAGAAA ATCCTACCAC CTTTCTGGAG AGACTAAGGG AGGCATTGAC
48701 AAAGCATATC TCTCTGTCAC CTGACTCTAT TGAAGGCCAA CTAATCTTAA
48751 AGGAAAAGTT TATCACTCAG TCAGCTGCAG ATATTAGAAA AAAACTTCCA
48801 AAGTCCGCCT TAGGCCCGGA GCAAAAGTTA GAAACCCTAC TGAACTTGGC
48851 AACCTCGGTT TTTTATAATA GAGATCAGGA GGAGCAGGCA GAATGGGACA
48901 AATGGGATAA AAAAAAAAAG GCCACTGCTT TAGTCATGGC CCTCAGGCAA
48951 GCGGACTTTG GAGGCTCTGG AAAAGGGAAA AGCTGGGCAA ATAGAATGCC
49001 TAATAGGGCT TGCTTCCAGT GCGGTCTCAA GGACACTTTA AAAAAGATTA
49051 TCCAAATAGA AATAAGCCAC TCCCTTGTCC ATGCCCCTTA TATCAAGGGA
49101 ATCACTGTAA GGCCCACTGC CCCAGGGGAC GTAGGTCCTC TGAGTCAGAA
49151 GCCACTAACC AGATGATCCA GCAGCAGGAC TGAGGGTGCC TGGGGCAAGC
49201 ACCAGCCCAT GCCATCACCC TCACAGAGCC CTGGGTATGC TTGACCATTG
49251 AGGGCCAGGA GGCTAACTGT CTCCTGGACA CTGGTGTGGC CTTCTCAGTC
49301 TTACTCTCCT GTCCCGGACA ACTGGCCTCC ATATCTGTCA CTATCCCAGG
49351 ACAGCCAGTC ACTAGATACT TCTCCCAGCC ACTAAGTTGT GACTGGGGAA
49401 CTTTACTGTT TTCACATGCT TTTCTAATTG TACCTGAAAG CCCCACTCCC
49451 TTGTTAGGGA GAGACATTCT AGCAAAAGCA GGGGCCATTA TACACCTGAA
49501 CATAGGAGAA GGAACACCCA TTTGTTGTCC CCTGCTGGAG GAAGGAATTA
49551 ATCCTGAAGT CTGGGCAACA GAAGGACAAT ACGGATGAGC AAAGAATGCC
49601 CATCTTGTTC AAGTTAAACT AAAGGATTCT GCCTCCTTTC CCTACCAAAG
49651 GCAGTACCCC CTTAGACCGG AGGCCCACCA AGGACTCCAA AAGATTGTTA
49701 AGGACCTAAA AGCCCAAGGC CTAGTAAAAG CATGCAGTAG CCCCTGCAGT
49751 ACTCCAACTT TACGAGTACA GAAACCCAAC AGACAGTGGA GGTTAGTGCA
49801 AGATCTCAGG ATTATCAATG AGGCCATTGT CCCTCTATAC CCAGCTGTAC
49851 CTAATCCTTA TATTCCGCTT TCCCAAATAC TAGAGGAAGC AAAGTGGTTT
49901 ACAGTCCTGG ACCTTAAGGA TGCCTTTTTC TGCACTCCTA TACATGCTGA
49951 CTCTCAATTC TTGTTTGCCT TTGAAGATCC TTCGAACCCA ACATCTCAAC
50001 TCACCTGGAC TGTTTTACCC CAAGGATTCA GGGATAGCCC CCATCTATTT
50051 GGCCAGGCAT TAGCCCAAGA CTTGAGCCAG TTCTCATACC TGGATATTCT
50101 TGTCCTTTGG TATGCGGATG ATTTACTTTT AGCCGCCCGT TCAGAAACCT
50151 TGTGCCATCA AGCCACCCAA GTGCTCTTAA ATTTCCTCGC CACCTGTGGC
50201 TACAAGGTTT CCAAACCAAA GGCTCAGCTC TGCTCACAGC AGAGGGCTAT
50251 TTATCCCTAA ATACTTAGGG CTAAAATTAT CCAAAGGCAC CAGGGCCCTC
50301 AGTGAGGAAT GTATCCAGCC TATACTGGCT TATCCTTATC CCAAAACCCT
50351 AAAACAACTA AGAAGGTTCC TTGGCATAAT AGGCATAACA GGCATAACAG
50401 GTTTCTGCTG AATATGGATT CCCAAGTACG GCAAAATAGC CAGACCATTA
50451 TATACACTAA TTAAGGAAAC TCAGAAAGCC AATACCCATT TAGTAAGATG
50501 GACACCTGAA GCAGAGGCAG CTTTCCAGGC CGTAAAGAAC ACCCTAACCC
50551 AAGCCCCAGT GTTAAGCTTG CCAGCGGGGC AAGACTTTTC TTTCTATGTC
50601 ACAGAAAAAA TAGGAATAGC TCTAGGAGTC CTTACACAGG TCCGAGGGAC
50651 CAGCTTGCAA CCCATGGCAT ACCTGAGTAA GGAAATTGAT GTAGTGGCAA
50701 AGGGTTGGCC TCATTGTTTA CGGGTAGTGG CGGCAGTAGC AGTCTTAGTA
50751 TCTGAAGCAG TTAAAATAAT ACAAGGAAGA GATCTTACTG TGTGAACCTC
50801 TCATGATGTG AACCGCATAC TCACTGCTAA AGAAGACTGG TGGCTGTCAG
50851 ACAACTGTTT GCTTAAATAT CAGGCTCTAT TACTTGAAGG GCCAGTGCTG
50901 TGACTGCGCA CTTGTGCAAC TCTTAACCCA GCGACATTTC TTCCAGACAA
50951 TGAAGAAAAG ATAGAACAGA ACTGTCAACA AGTAATTGCT CAAACCTACG

FIGURE 3-15

51001 CCGCTTGAGG GGACCTTCTA GTGGTTCCCT TGACTGATCC CAACCTCAAC
51051 TTGTATACTG ATGAAAGTTC CTTTGTAGAA AAAGGACTTC GAAAAGCAGA
51101 GTGTGTAGTG GTCAGTGATA ATGGAATACT TGAAAGTAAT CCTCTGACTC
51151 CAGGAACTAG TGCTCAGCTG GCAGAACTAA TAGCCCTCAC TCAGGCACTA
51201 GAATTAGGAG AAGGAAAAAG GGCAAATATA TATACAGACT CTAAGTATGC
51251 TTACCTAGTC CTCCATGCTC ACGCAGCAAT ATGGAGAGAA AGGGAATTCC
51301 TAACTTCTGA GGGAACCCCT ATCAAACATC AGGAAGCCAT TAAGAAACTA
51351 TTATTGGCTG CACAGAAATC TAAAGAAGTG GCAGTCTTAC ACTGCTGTAA
51401 GAAAGGACAG AGAAATAAAA GGGAACCGCC GAGTGGATAT TGAAGCCGAA
51451 AGAGCCACAA GGCGGGACCC TCCATTAGAA ATGCTTATAG AAGAACCGCT
51501 AGTATGGGGT AATCCCCTCC AAGAAACCAA GCCCCAGTAC TCAGAAGAAG
51551 AAATAGAATG GGAAACCTCA TGAGGACGTA GTTTCCTCCT CAGGATGGCT
51601 AGCCACCAAA GAAGGAAAAA TACTTTTGCC TGCAGCTAAC CAATGGAAAT
51651 TACTTAAAAC CCTTCACTTA GGCATTGATA GCACCCATCA GATGGCCAAA
51701 TCATTATTTA CTGGACCAGG CCTTTTCAAA ACTATGAAGC AGATAGTCAG
51751 AGCCTGTGAA GTGTGCCAAA AAATAATCCC CTGCACTTCA GGCCATGCAT
51801 TTCAATCCCT GAATCTTTAA CCTCCTTGTT AAGTTTGTCT CTTACAGAAT
51851 TGAAGCTGTA AAGCTACAAA TGGTTCTTCA AATGGATCCC CAGATGCAGT
51901 CTATGACTCA AATCTACCGC GGACCCTTGG ACCGGCCTGC TAGTCCATGC
51951 TTCGATGTTG ATGATATCAA AGGCACCCCT CCCGAGGAAA TCTCAAGTGC
52001 ATGACCCTTA GTTGCACCAG TTCAGCAGGA AGCAGTTAGA GCGGCCGTTG
52051 GCCAACCTCC CCAATAGTAC TTGGGTTTTC CTGTTGAGAG GGGTTGCTGA
52101 GAGACAGGAC TAGCTGGATT TCCTAGGCCG ACTAAGAATC CCTAAGCCTA
52151 GCTGGGAAGG TGACTGCATC CACCTTTAAA CACGGGGCTT GCAACGTAGC
52201 TCACACCCGA CCAATGAGGT AGTAAAGAGA GCTCACTAAA ATGCTAATTA
52251 GGCAAAAACA GGAAGTAAAG AAATAGCCAA TCATCTATCA CCTGAGAGCA
52301 CAGGGGGAGG GACAATGATC AGGATATAAA CCCAGGCGTT CTAGCCGGCA
52351 ACGGCTACCC TCTTTGGGTA CCCTCCCTTT GTATGGGAGC TCTGTTTTCA
52401 CTCTATTAAA TCTTGCAACT GCACAAAAAC CAAACCAAAC CAAACCAAAC
52451 AAACAAAAAA ACAGTGACTG ACTTATGGTA AACATTATAT AAGCATAAAG
52501 TAAACCAAAT ACTTTTTTTC TAATTATAAA AGTCCTACAC TAACATTGCA
52551 GAAAACTTGA GGAATTCAGA AAAGTTATTA CTTAGTAAGA GTTGGAATGA
52601 ATAAATAAGT GGGTAGTTAG GATGGCAGGC ATGTGTTTTA GGCAGAGAGA
52651 TACAAGATAA AGAACTAAAA CTAGAATCTG GTCTTTGAAC CCCTGGCCTG
52701 ATTGTCTTAT TCATCATGAT GATTTGCCTA TTTTTCCAAT TTCTAAATCA
52751 TTCCTCTGCT GTTGACAAAG CAATAAATTG TTATATTTGA TAAGTGAATC
52801 TTCAGAGAAC TGGCCTTGAG CCAGCTCTAC AACTAACCAG CTCTGTGGCC
52851 CTTTGGAGAA TTTCTTAATA TTTGTAAACC TCAGCTTTCC TACCAGTGAA
52901 ATGAAGTTAG TCCTCCCTGT CCTGCAGGGT TGCTGCAAGG ATTTAACAAC
52951 ATGTATATGT ACAAACCACT TAGTCCTGTG CTTGGCCTAT TTGGTGCTTT
53001 TTTTTTTTCT TTTTTTTAAG ACAGGGTCTT GCTTGAATCT TGCTGAGGCT
53051 GGATTCAAAC TCCGGGGCTC AAGTGATCCT CCTGCCTCAG CCTTCCAAGT
53101 AGCTGGGACT ACAGGCCTGC ACCACTGTGC CTGGTGGCAG TGCTCGTTGA
53151 ATGTTCTTTT TCCTTAGTTC CTTCCTAGCT CTTCTGACAG TTTTGGGGCT
53201 TATGTATATA AGAAGGACTT GGTTGCCTCA GGGAGAGAGG ATGCAGTAGA
53251 GTTACATAGC TCACCTCACA TCCTCCAAAA GCTGAATTCA TAAGTAAACA
53301 AAGTGAGCAT TTCACCCATA CTTTACACAA AGTCTAGAAT ATTTATGGTG
53351 TCCATCAGGC TCACATACTG TGACCTTCTG AGATACTTTT CCCTCTCCAT
53401 TCCCTTTCCT TCTCCCTGCT GGCTTTTTTT TTTCTTCTTC TTCTTTTTTT
53451 TTTTTTTTTA CTGTGAAAAA CAACCTATAT ACAGAATAGT ACAAAAACAT
53501 ACCTGTATAG TTTGAAGAGT AATTATTAAC AGTCTTATTA AGAAACAATG
53551 CTCCATCCAT GTTACTGCAA AAGACATGAC CTTATTCTTT TTCTTCTTAA
53601 TTTTTTTCTT TTTCTTTCTT TATTTTGGCC CTTTTTCAGA TCTAGACCTG
53651 CAGAGATCTT GTTCTTTTTT TGAGACAGCA TCTCGCTCTG TCACCAGGCT
53701 GGAGTTCAGT GGCGTGATCT CGGCTCACTG CACCCTCTGC CTCCTGGGTT
53751 CAAGTGATTC CCCTGCCTCA GCCTCCTGAG TAGCTGGGAC TACAGGCGCG
53801 TGCCACCACA TCCAGCTAAT TTTTTTATTT TTAGTAGAGA CTGGGTTTCA
53851 TCATGTTGGC CAGGATGGTC TCAATCTCTT GACCTGGTGA TCTGCCTGCC
53901 TCGGCCCCCA AAAGTGCTGG AATTACAGGC ATGAGCCACC ACGCCCAGCC
53951 GATCTTGTTC TTTCTTATGA CTGTGTAGTA TTCCATGGTA TATATGTACC
54001 ACATTTTCTT TATGCATTCT ATCATTGATG GGCATCTAGG TTGATTCCAT
54051 GTCTGCTATT GTGAACAGTG CTGCAGTGAA CATTCACGTG CATGTGTCTT
54101 TGTGGTAGAC TGCTTTATAT TCCTCTGGGT ATATGCTCAG TAATAGGATT
54151 GCTGGATTGA ATGGTAGTTC TTCTTTTAGC TCTTTGAGGA TACTGCTTTC
54201 CACAATGGTT GAACTAACTT ACACTCATAC AGTATATAAG CATTCCTTTT
54251 TCTCTACAAC CTCGCCAGCA TCTGTTACTT TTTGACTTAT AGGTGGGAGC
54301 TAAATGATAA GAACTTATGA ATGTAAAGAA GGAAACAGAC ACTGGGATCT
54351 ACTTGAGTGG GAAGAGTGGG AAGAGGGAGA GGAGCAAAAA AGATAACTCT

FIGURE 3-16

54401 TGGTTACTAA GCTTAATACT TGGACAATGT AATAATATGT ACAACAAACC
54451 CCTGTGACAC ATGTTTATCT ATGTAACAAA CCTTCACATG TACCCCTAAA
54501 CCTAATTTTT TTTAAAAGAA ACAGAATGCC AGCCAGGCAT AGAGGCTAAT
54551 TGCCTCTAAT CCCAGCACTT TGGGAGGGTG AGATGGGCAG ATCACTTGAG
54601 CCCAGGAGTT TAAGGCCAGC CCAGGCAACA TAGCAGAACC CCATCTCTTC
54651 AAAAAGTACA AAAATTAGCT GGGCATGGTG GTGTGCACCT GTAGTCCCAG
54701 CCCCTTGGGA GACTGAGCTG GGAGGATGGC TTGAGCCCAG GAGGTCAAGG
54751 CTGCAGTGAG CTGTGATCAT GCCACTGCAC TCCAGCCTGG GCGACACAGC
54801 AAGACCCTGT CTCCAAAAAA AAAAAGAAAA GGAACAGAAT GTTACCAGCC
54851 CAACTGCACT TCTTCACTCC CCCACCACTC TAATGAAGTC ATCACTAACC
54901 CACTTCTAAA GTACTCACAT ACCCTATGTC TATGGAGGTA TGTCAGTGGA
54951 GGCTAAGGTA TGCCAGTGGA GGCTTATGTA CCTTATGTCT AAATATTTAA
55001 AGTTATTAAA TTAAAAAACC ATTAAAATAT GCTTTCTACC TTGACAAACC
55051 TTTATAACAA AATTAGAAAA TGTTTAATGT TATGGCATTA AATAATTGAA
55101 AGCAAAATAT CAAAGATGAT AGAATTTAAT TAATTATTTT ATTTTATTTT
55151 ATTTGAGAGA GGGTCTTTCT GTGTCACCAA GGCTGGAGTG CAGTGATGCA
55201 ATCATGGTTC ACTGCAACCT CAACTTCCCG GGCTCCAGTG ATCCTCCCGC
55251 CTCAGCCTCC CAAGTGGTTG GGACTACAGA CATGTGCCAC CAAATCCAGC
55301 TAATTTTTAA ATTGTTTTTA ATAGAGGTAA GGGTCTCACT ATGTTGCCTA
55351 GGCCAGTCTC GAATTCCAGG GGCTCAAGGG ATCCTTTTGC CTTGTCCTCC
55401 CAGAGTGCTC GGATTTAAGT TGGGAGCCAC TATACCCACC CAACATAATT
55451 CAATTATTTA ATATTTCACA TGTTTTAGTA TTCCTTTGAT AGGGATGTGA
55501 TGTTTGGGTG AATAATAAAG TAAATCAAAG ACATATATTT GAAAATTATG
55551 TAGTTATTCT AAAAAATTAA TTATTTACCT TTATTTTAGC AAAATCAGTG
55601 TGTTAGCATA ATCAAGATAT TTTGGTATTC TAGTAACAAG ATCTAGTCAC
55651 AGTAATGATG TAAAGATTAA AAAATAAAAT ATAATAGGAA CCAGTATAAA
55701 CAAGTGGAAT TTAAATTTAA AATGCAATAC CAGCTGGGTG CGATGCCTCA
55751 CGCCTGTAAT CCCAGCACTT TGGGAGGCCA AGGCAGGCGG ATCACCTGGG
55801 GTCAGGAGTT CAAGACCAGC CTGACCAACA TGGAAAAACC CCATCTCTCC
55851 CAAAAATACA AAATAAGTTG GGTGTGGTGG TGCATGCCTG TAATCCCAGC
55901 TACTCGGGAG GCGGAGGCAG GAGAATCACT TGAACCAGGG AGGCAGAGGT
55951 TGTGATGAGC CGAGATCACA CCATTGCACT CCAGTTTGGG CAACAAGGGC
56001 AAAACTCTGT CTCAAAAACA AAAAGAAACA AAAAACACAG TACCATTTAC
56051 ATTAGCACCC CTCAAAATGA AATACTTAGG TATAAATCCA GCAAAATAGG
56101 TATAAGAGAT ATATAAGTAA AACTATAAAG CTCTGATGAA AGAAATAAAA
56151 GAACCAAATA AATGGACAGA TATTCCATGT TCATGGATAG GAAGACTCAG
56201 TAATGTCAAG ATGTCAGTCC TTTTCATCTT TATCTATAGA TTCAGTACAC
56251 TTCCAATCAA AATCCCCGCT AGTTATTTTG TGGATATTGA CAAACTGATT
56301 CTAAAGTTTA TGTGGACAGG CAAAAGACCC AAAATAGGTG ACATGCTATT
56351 GAAGCAGAAT AAAGTTAAAG TCTCTACTTC TTCTTCAATA TCATTTTCTT
56401 CAATATCAAA TCAACTCATT TTCTTCAATA TCAAGTCAAC TGTAGACTGA
56451 CAGTGTCCTA CTTCAAGACT TACTGTAAAG CTACAGTAAT CAGAACAGTA
56501 TGGTATTGGT GAAAGAATAG ACAAACAGAT CGATGGAACA GAATAGAGAG
56551 CCCAGAAATA GACCCAGACA AATACAGTTA ATCTTTGACA AAGGAGCAAA
56601 GGCAATAGAA TGGAGGAAAG ATACCCTTTT CAACAAATGC TGTTGGAATA
56651 ACTGGATGTC CACATGTAAA AAAATGAATC TAGACACAGA CCTTACACCC
56701 TTCACAAAAA TGAACACAAA ATGGATCATC AACCTAGACA TGAAACACAA
56751 AACTATACAA ATTTCTGCTC TGTGAAAGAC AATGTCAAGA GAATGAGAAG
56801 ACTTGGAAAA TATTTGCAAA AGACACATCT GATAAATGAT TCTTATCTAA
56851 AATATATAGG AACTCTTAAA ACTGAACAAT AAAAAAGAAA CCTGATTTTA
56901 AAATGTGCCA AACACTCTAA TAGATATCTC ATCAAAGAAG ATATATAGAT
56951 GGCAAATAAG CATATGCAAA GATGCTCCAC ATCATAGGTC ATCAGAAAAA
57001 TGCAAATTAA AATGATGAGA TTCTACTACA CACTTATTAG AATGGCCAAA
57051 ATCCAAAACA CTGACAACAC CAAGTGCTGG TGAGGATGTG GGACAACAGG
57101 AACTTTCATT TATTGCTGGT GGGAATGCAA AATGGTATAG CCACTTTGGA
57151 AGACAGTTTG CAGTTTCTTA CAAAATTAAA CACACTCATC ATATGATCCA
57201 ACAATTGCAG TCCTTGATAT TTACCCAAAG GAGTTGAAAA CTTATATTCA
57251 CACAGAAACC TGCACATAGT TGCTTACAGC AGCTTTATTC ATAATTGCCA
57301 AAACTTTGAA GCAACCAAGA TGTCCTTCAA TAGGTAAATG AATAAATAAA
57351 CTGTAGCATA TCCAGAAATG GAGTACTCAG TGCTAAAAAG GAATGAGGTA
57401 TCAAGCTATG AAAAGACATG GAGGAACCTT TATATTTTTA TTTTTTTAGA
57451 GACAGGGTCT TGCTCTGTCA TCCAGGCTGG AATGCAGTGG CACAATTATG
57501 GCTCACTGTT GCCTTGACCT CCTGGGCTTG AGCTCTCTTC CTGCCTCAGC
57551 CTCCCAAGTA GCTGGGACTA CAGGTGCATG TCACCACACC TGGCTAATTT
57601 TTTTTTTTGA GAGATTGGGT CTTGCTGTGT TGCCCAGGCT GGTTTTGAAC
57651 TCCTGGGCTC AAGTGATCTT CCTGCCTCAG CCTCCCGAAG TGCTGGGATT
57701 ATAAGTGTGA GCCACTATGC CTGACTTTTT TTTAAATTTA TTTTTCTTCT
57751 AGAGACAAGG CCTTGCTTTT TATTGCCCAG GCTGGAGTGT AGTGATGCAG

FIGURE 3-17

57801 TCATAGCTCA CTGTGGCCTC AAGATCCTGG GCTCAAGTAG GAGCCCAGCT
57851 AATTTTTTTT AAAATTTTTG TAGAGATGGA GCCTTGCAAT GTGGCCCATG
57901 CTATGGAGGA AACTTAAATG CATATTTCTA AGTGAAGAAG CCAGTCTGAA
57951 AATGTTATAT ACCATGGGAT TTCAACTATA AGACATTCTG GAAATGGCAA
58001 AACGAAGGCA ACAATAAAAA GATGAATTGT CAGGGAGTTG GTCAGGGGAG
58051 AATGAATAGG TGGAGTACAG AAGATTTTTA GGGCAGTGAA ATGTCTCTGA
58101 TACAGTAATG GTGGACACAC GTTGTATATT TGTCCAAATC CATAGATTCT
58151 ATAATACCAA CAGTGAACCT TAATATAAAC TATGGACTTT GGGGCTGGGC
58201 ATGGTGACTC ATGTCTGTAA TCCCAACACT TTGGGAGGTC AAGATGGGAG
58251 GATCACTTGA GGCCAGGAGT TTGAAACCAG CCTGGTCAGC ATTGTGAGGC
58301 CCTATCTCTA CAAAAATAAG GAAACCATGG ACTTTGGATT ATAATGATGT
58351 TTCAGTGTAG GTTCCTCAGT TATAACAAAT GTACTACTCT GGTGGGGGAT
58401 GTTTATAATA AAGGAGGCAA TGCATGTGTT GGGGCAGGAG GTATACGAGA
58451 AATCTGTTTA CCTTCCTCTA AATTTTACTG TGAATCCAAA ACTGCTCTTA
58501 AAAAAAAAGG TCTTAAAAAA TAAAATTATA TTTGAGGGAA AATATTTGAA
58551 TTATTATTAT TATTTTCTTT TTGAGACGGA GTCTCTCTCT GTCCCCCAGG
58601 CTGGAGTGCA GTGGCACGAT TTCGGCTCAC TGCAAGCTCT GCCTCCCGGG
58651 TTCACGCCAT TCTCCTGCCT CAGCCTCCCT AGTAGCTGGG ACTACAGGCG
58701 CACGCTGCTA TGCCCGGCTA ATTTTTTGTA TTTTTAGTAG AGATGGGGTT
58751 TCACCGTGTT AGCCAGGATG GTCTCGATCT CCTGACCTCG TGATCTGCCC
58801 GCCTCAGCCT CCCAAAGTGC TGGGATTACA GGCATGAGCC ACCGCGCCTG
58851 GCCTATTTGA ATGATTTTTA TCAAAGATGT AAATTAAAAC AAATGTAAAA
58901 ATAAAAAACA AATCACTGTC TGATTCTATT TGTATAAATG TCTAGAAAAT
58951 GCAAACTAAC TTATAATGGC AAACTACTCT ATAGATCAGC AGTTGCCTGG
59001 AGGCAAGAGG GAAGAATTGC AGTGAGGTAT GATAAAACTT TTGGGGTAAT
59051 AAATATATTC ATTATCTTGA TTGCAGTGAT GTTGGCACAG GTACATCCAT
59101 ATGTCAAGAT TTCTTGTTGA ATACTTTATG TAGTTTATTG CATAACAATT
59151 CTATAAAATT AAAAATCATA AAATTTTGTT TGTTTTAAAA ACATTCTTTC
59201 TTTCTTTTTT CCTGAGACGG AGTCTCCCTG TATCACCTAG ACTGGAGTGC
59251 TGATTGCAGC CTTGACCTCC TAGGCTCAAG TGATCCTCCA GCCTCAGCTT
59301 CCTAAGTAGC TGGTACCACA CAGGCGCATG CCACCACACC CATATAATTT
59351 TTAAATTAGT TTTTGTAGAA ATAGGGTCTT GCCATGTTGC CCAGGCTGGT
59401 CTTGAACTCC TGGGCTCAAG CAATCCTCCA TCTTGGCCTC CCAAAGTGCT
59451 GAAATTACAG GTGTGAGCCA CTGCACCTGG CCATCTTAAT TTTTAATATT
59501 TAAAAGAAAA GTAAGGGCCA GACACTGTGG CTCTCACCTG TAATCCCAGC
59551 ATAAGACCAG CCTGGGTAAC ATGGCAAGAC CCCATCTCTA TCAAAAATTG
59601 AAAAATTAAC TGGGCATGGT GGTGGCCTGT GGTCCCAGCT ACTCAGGAGG
59651 CTGTAGCTGG AGGATCACTT GAGCCTAGCA CGTTGAGGCT GCAATAAGCC
59701 ATGTTTGCAT CACTGCACTC CAGCTTAGAT GAGAGAGTGA GACCCTGTCT
59751 CAAAATAAAT AGATAGATAA TATATGTCGT AGTTTTAAAA ATATATTATT
59801 AAGATAAAAA GCAAGCCAAG ACAATTAAGT GGGGGAAGAA TAGTTTTTCC
59851 AACAAATGGT GCTGGAACAA CTGCATAGCC ACAGGAAAAA GAATGAAGTT
59901 AGATCCCCTA CCTCACACCA TATAAAAAAA TAACTCAAAA TGGATTAAAG
59951 ACCTAAGTAT AAGCTGAGAC AAGAAGATTA CTTGAGGCTT GGAGTTCAAG
60001 ACCAGCCTGG GCAACATATT GAGACCTCGT CTCTTAAAAA AGAAAAAAAA
60051 TCAGCCGGGC ATGGCAGTGT GTACCTGTAT TCCTAGCTAC TCAGAAGGCT
60101 GAGGCCAGAG GATTGCTTGA GCCCAGGATT TAGAGGCTGC AGTGAGCTAT
60151 GATTGCACCA CTGTACTCCA GCCTGGGTGA CAGAGTGAGA CCTTGTCTGC
60201 TCCACCCCTC CTCCACAAAG TGTAAAGGTA TAAATGTTTG TGGCCTTGGA
60251 TTAGGCAATG GTTTATTAAA TATGACATTA AAAGCACAAA CAACAACAAA
60301 ATAGATTAAT TGGACTTCAT CAAAATTAAA ACCTCTGTGC TTCAAAGGGC
60351 ACACCAAGAA AGTGAAAAGA GAATCCACAC AATGGGAGAT AATTTTTTGC
60401 AAATCATGTA TTTTACAAGA CTGGTGTCCA GAATATATAA AGAACACTTG
60451 CAACTCAGCA ATAAAAAGAC AAGTAACACA ATTTAAAAAT GTTGAAAGGA
60501 TTTGAATAGA CATTTCTTCA AAGAAGACAT ATAAATCACC AATGAGCATA
60551 TGAAAATGTA CTCAACCTCA TTGGTCATTA GAGAAATGCA AATAGAAGTC
60601 ACACCCATTA GGATGGCTAA AATAAAAAAA GATGAACAAT AACAAATGTT
60651 GGCAAGTATG TGGAAAAATT AGAACCCTCA TACACTGTGG ATGGGAATGT
60701 AAAATGGTGC AGACACTTTG GAAAGTTGGC TATTCCTCAG AGATTTACCA
60751 CATGGCACAG CAATTCTACT TTTAGGTGTA TACCCAAGAC AATTAAAAAG
60801 ATATATACAG GCCGGGCGCG GTGGCTCAAG CCTGTAATCC CAGCACTTTG
60851 GCCAAGGTGG GTGGATCACG AGGTCAGGAG ATCGAGACCA TCCTGGCTAA
60901 TACAGTGAAA CCCCATCTCT ACTAAAAATA CAAAAAATTA GCTGGGCGTG
60951 GTGGGGGGGC GCCTGTAGTC CCAGCTACTC GGGAGGCTGA GGCAGGAGAA
61001 TGTCGTGAAC CCGGCAGGCG GGGCTTGCAG TGAGCCGAGA TTGCGCCACT
61051 GCACTCCAGC CTGGGCAACA GAGCGAGACT CCGTCTCAAA AAAAAAAAAG
61101 ATATATACAC AGAAAAACTT GTACATAGAT GTTCATAGTA GTATTCCAAT
61151 AAACATGCCC ATCAGTAGAT GAATGGATAA GCAAAACGTG GTGTATTAAT

FIGURE 3-18

61201 AAATGAAATA TTATCCAGCC ACAAAAAGCA ATGAAGTACT GATACATGAT
61251 CCAATATGGA TGGACCTTGA AAACTATACT AAATGAAAGA AACCAGCCAC
61301 AAAAGGCCAC ATAGTACATG ATTACATTTG TATAAAATGT CCAGAATTAG
61351 CAATTCCATA AAGACAGAAA GTAGATTAGT AGTTGCCAAG GGCTGAGGGA
61401 AGGAGGAATG GGAGTGACTG CTAATGGGTA CAGGGTTTCT TTTTGGGGTG
61451 AGAAAAGTGT TCTGGAATTA CATAATGATG ATAGTTGTAC AACTTTGTGA
61501 ATATACTAAG ACACACTGAA TTGTATCTTT TAAAAAGTTA AATTTTATGG
61551 TATGTGAATG ATACCTCATT AAAATAGTTA CATGAGAAAA AAATCAAAGG
61601 CAAAATACAG AGTATAATTC AAGTATTTTA ATTTTTAAAT ATAAAGTATT
61651 TATAGCCAAA TTTGATTTAC TTCTAAAATG TCTTATTAAA TAGTTTAATA
61701 AAAGCAAAAC TGTTCTAGCA TTCAGTGTTT ATTAATTTGC AATACACAAA
61751 CAATATCATG TTTTACTCAT GTTGGGTCCA CCACATATGT ATATATTTAA
61801 GAATAATGTG ATTGGTCAGT ACTGCAAAAT GTTTTCGTGT ACGGTGGCTG
61851 TGAGTACCAT ACTAATTAGT ACACTAAGAA CTATGAATTG GAACAGGAAG
61901 AAAAGCAAGA AAATGAACAT TCAGCACTAT TTGGAAATGA AACCTCACTA
61951 TGCAAGAATT CATTGCATTC ATGCCTTCTC AGGGGGAGTG TTTGATAAAT
62001 TAATATTTCA CTAACCTAAG AGCTTCTGCT GCTCAAATCC TCTACACACA
62051 CTAATGCAGT GTCAGTCTCT GAGTTTGGCA GTGGCACAGC CATAACTTTA
62101 TGGCAGTTAG GTGAAATCCC TTTTTGTTTT GTGGACATAG AACAACAGAT
62151 GTGGAGTTGT ACTTCTCTGG GGCCAGCACA AACGCCATTA AGTCTGTGAG
62201 TCTCATTCTG TCAGTGCTGA GCATCCATGT GTTCTTTGAT TTGCGAGTGT
62251 GTGTTTGTGA TTGGCAGGTT TCTCTTACGT GACTCTGAAG GAGTGTTGCT
62301 GATCACATCA CACCCATCCT TAGGCCTTAT TGATCATCAG TGTACCTTCC
62351 CACTACTATA CTTTAAATAG ATGCCTGTTA TTTAAATTTG ATTTTGAGAT
62401 TAAACAGAAT GGCAGAGACA ATTTTGAGAC ACATCTTTCT TTCATGCTCT
62451 GGTAGGAAGA TCAAGATTTC TAGGACAGTA GAACAGAGTA AAAGATGAGT
62501 GCTGTTGGGA TCCTATCTTT CTCCTAAGCT ATTTTTTCTC TCTCAGTTAT
62551 TCATCATCTA TCTCAATTTT TCCTAATGAA CTCTTCTATA TAAAAGAGGA
62601 TCCAGGTCCC ACATCCACTG TCAAGGAGGA ATGTAAGATT GACTTGCAAC
62651 TCAGCCTGTG TACGCAGTTT ATGGTTTTTC TTGGCAGGTT TAATGTTCTT
62701 TCTTATCTCT TAACTTCTTG CTATTCAATA GTAAGTAACT CCCTGGCAGA
62751 ATTACCTGTG GCTAGAGAAT GCTGTTATCA GAGCATCTTT GTTTAATTGG
62801 TACTTAGAAC AGAAGGTGTC ACCTATTTGA CAGGCCAACA ATTATGAGCA
62851 AGGAGGCATT TGATTCATCA AGATAGAAAT CTGCCTGTTA GGTGGAAACA
62901 TGTCTATGTG GGTTGATATG TTTTTAGAAT ATTAAGGCTT GTTTGTGCAT
62951 GACAACTTTA GGAAGGTGTA CTCCAAATGT CTCCAAAGGT TTGCTGTAGT
63001 TCTTTACAGA AGTTGGGCTG CTCCTGGTGG ACAGTGTGTA ACAGTGAACA
63051 ATGTATGCTC TAGACTGGGT TCCCTTCCTC CACCCTGTGT CTGTGTGGCC
63101 TTGGGCAAGT TGTTTAACCA ACCACTTTTT GCCTCAGTTT CTTTATCGGA
63151 ACAAGGAGAA TAAGAATACT TCAATCAGGC CAGGCGTGGT GACTCACGCC
63201 TGTAATCCCA GCCTTTGGGA GGCCGAGGTG GGTGGATCAC CTGAGGTCAG
63251 GAGTTCCAGA CCAGCCTGGC CAACATGGTG AAACCCCATC TCTCCTTTAC
63301 TTATGCTGGC CTGATATTGA TCGTCCATGG TAGAATTGAT ACTGCTTGAC
63351 AAAGCAGCCT ATTTCAGTCA GGACCCTTCT TCTCTAGTTT CCTCTGTAGC
63401 TATTACCTTA GCCTTCCATT TCATTCTTCA CACTACAGAT ACTCTCATTG
63451 ATAAAGGAAT GATGTCTTTA TGCTTTCAAG CATTCTGGCA AGTTAGTAAT
63501 TCAACTATGA TTCTAGGTCA GACAAAACCA GTTATGAACA TAAGACTGTT
63551 TTTAATCTCC TCCCTGGTCC CCCAACCACC CACCCCAATC AGGAGAAACT
63601 ATGTTCTGCA TTGGTTTAAG GAACCCGCTT CTTTCTTTGA TACCTGACCT
63651 ACAGATCCAA TCTATTCCCA GGAATTTTGA TAAGAATTCT CAAATCCTCA
63701 GCAAGGCTAT GCCACTGTCA TGACTCTCCT ATTCCTGGTA GTGTCATTCT
63751 CAGTGTAGGC TGTTTGATAG GTAGTTTTGT GAAGTCTTGT TCATCATAAT
63801 GGATCATATG ATTTTTAAAA GCAGGACCTG GGTCATATGC CCAGATTAAT
63851 TTCAACAAAG TTGGTATGTT TTCTTCCTAA AATTAATTTT TTATGATTAT
63901 CAAAGTTTTA TATGCATACA GTATAAAATA TCAGATTTCT ATAAGGCTTA
63951 TAATAAAGAG TAGCATTCCT CTGCCCTATT CCCCTCTCAT ACTGGGCACC
64001 TTACTCCAAA ATCAACTGCT TTCAGCTCTT TTAGCCATTT CTTATGGTTA
64051 TCTTCATATT TCAAAACAGC ATGCTTATAT GGTTGTATTT TGAATTTTCA
64101 AACTTAGATT TTTATCTACA GACTTCCTTA TGGAAGATAA ATATTTAACT
64151 CTTTTTGTGC TACTCCTTTT CTCATCTCTT AATGTAGGCT ATGTTAAAAT
64201 TTTTGGTTAA ATCAATTAAA AGCCAGTATA GTATAGTGGC TTAAGAGTGA
64251 GGGTGCTCAC CTCAAATTCC AGCCTCACCA CCCATTACTT TGTGTGACTT
64301 TGAGCAAGCT TTAAAACGTC AGTGTCTCAG TTTTGTCAAC TGAGTAGATA
64351 CCTCATAGAA TTGCTGTTGA TATTAAGTGA CTTAATCCTA TGGGCTGAAT
64401 TTGTCTCCCA AAGTTCATTT GTTGGAAACT TAATCCCCAG TGCATGTGTT
64451 AAGAAGTGGG ACCTTTCAGA GTTGAATAGG CTATGAGGGC TCTGCCTTCA
64501 TGAATGGATT AATGCCGTTG TTGCAGTAGT GGGTTCCTTA TAAAAGGAAG
64551 TGTTTGACCC CCTTTCCTTG CCTCCTCATG CATGTGATGG CCTTAGCCAT

FIGURE 3-19

64601 GTTATAATGC AGTAGTAACG CCCTTACCAG ACACTGGCTC CTTGATCTTG
64651 GACTTCTCAG CCTCCAGAAC TGTAAGAAAT AAAAGTTTTT TCTTTATAAA
64701 TTACCCAGTC TCTGGTATTC TGTTATGGCA GCAAAAAACA GACTGAGACA
64751 CTTAATATAT ATGAAGCATC TAGACTGTCT GGCACATTGT ACATTTTAAA
64801 TCCCAGATAT CGATATCATC AATATCATCA TCATCATCAT CTGTGGCTGT
64851 ATAATACCTC CCTCTGCATT TAAAGGATGA GGGCTGGTGT AGCAGTTATT
64901 AATATAAGTG AGCTAGTTGG CTATAAACCT CTCCTATAGG CTTTGCCATA
64951 AACTTGTGTC ATGGATAATC AGAGAATTTG GACCTCCTAA TGACAGGCCA
65001 CTGACAAGAA AAGCCCAGAG GGAGCTGATT GAGCATGCTC AGTTCTTTCT
65051 ACCAAAGGCC TCAATCAGAC AGATTCTCTT TCCCAGGAGT AGACACTGAG
65101 GGGGTGGAAA CAGGCTCTGA GTTGACTCCA CTGAGAAGTC CCTGAATGGA
65151 TAGCAGCCAG GGAATTAAGG AGTTTCCCTG TGGCAAGTCT CTACTCGTAC
65201 AATATTAGGA CAGCTGTTTT TTTAATTTGT TCTGTGGGTG CATATTTTTG
65251 ATCTCCACAA CATAACTACT TACTCACTGT GTTGCTTTTT GAAATTTATA
65301 ACATCTAAAT TTGCAAATGT GAGGTATGCT CTCAGGAAAA TTATATATCT
65351 GTTAAAATTC TTTGCTCCTT TTAGGACACT TCCATCCATC TGGTGACCTC
65401 TAGTAGCGTA CATACTCAGC AACAATTAAA GGTAGTTTCT GGCTTGTTTT
65451 TCCCAAACAA TAATTTGTTC AAAATACAGA AATTTGGAAA GTACCCCATG
65501 ATGAGAAACT TTTGAAAGAG AATTTAATGT ACAGTAACTA TTTTTTTCCT
65551 CTGAGGAATA ATTTTGGGAA AAGAAATTTG CTTTATATTC AGGCATATTG
65601 AGTAAGTTGC CTTCATTTTC AAACAAATTC ATCACCTCTC CTTTCCCCAT
65651 TCCCCAAACT TGTTCATTTT CTTCATTCTT AAACTCTGCA TCTGTCTTCC
65701 CCTCACTTTA CCACGTAGCT AAATCTCTTA AGAGCCAACC TATTGCCACT
65751 CCCTTCATCT TTTTTAATCT AATGGATCAT AAGTCCTGTT GATTTCATCA
65801 CTCACATCTC TCAAATCGGT CTCCTTGGAT ATCTTGCTAC TTCTGATTTC
65851 AGACAATTGT CATTTGTCAC TTGCAGCAGC CTTCCAGCTG CCTTCCTCCC
65901 AATCTCCCCT ATGAGCATTC TGTTCTTCAA ACTGTTGCCA AAGAGGAATT
65951 TCTAAAATAC CGTTTGATCA TGTAATTCCA GGTTTTAAAA ACCTAATGTT
66001 GACTCTCTGA CATTTAAAGA ATAAAGTCCG TATTTCTTGG GATGACTGAC
66051 ATGGCTTTGC TGTGATTCTG GCTACATCCT TAGAGAATCA CTTGCAGATA
66101 TTAATTTTGC ATATGCTATT CCCTTTACAT GGAATGTTCC TGCCTCCTCT
66151 TTCCCCCATC CTTTGCCTGA TGAACGCCTA TTCATCCTCC ATGTCACTTT
66201 GGGTGTTACC CTGAGGAAAC AGTGGTTATT TTTTCCCTGC CCATCTGGGC
66251 TGGATACCCC TCCTAAGTTG TCCCACAACA CACAGTACAC ACCTTGGTCG
66301 ATTATCACAC TGTTTGTTGT TCTGCCGTTG TTTGTGTCTC TTGAATATGA
66351 GTTTCTTGAG ATTAGGGACC TGAGATCCCC AGTGCCCAGC AGAGCAGGAC
66401 CTGTTTCCAC CCCCTCAGTA ACTACTCCTG GGGAAGAGAA CCTACAAATA
66451 AATGTGTATT GCATGAATAA AGCTATATAT CCCCTTCCTG TTCATACTTA
66501 TCCATTTAAT TTCTGAACTC TAAATGCTGT ATTTTTCTCC ATTAAATTTG
66551 GTTTTATTGG TTTCAGTTTC CTCTTTATAT GGATTTTGAG TCCTAATTTT
66601 GTCAACCAGC ATATCAACAG TTCTTTCTAG CTTTTGTCCC CCGGTAGAAT
66651 GGCAGCTCCA TGAAGACAGG GATTAGTGTC TGTTTTGTTC ACTGCTGTTT
66701 TCTCAGCATC TAGAATAGTG CTGGCACATA TAGATACTCA CAAAGTATTT
66751 GTTGAATGAA TGAAGGTGTT GCATACAAAT TTGATAAATA AAACTTAGGT
66801 TTTCTTTCAA ATATTAATCC TTAGCTCCAT TCTGTTCAAT ATTTTTATTA
66851 GAGATAACCT TTAAAAATCT TCCTTGTCAC ATAAGAGTTA ATTCCATGAA
66901 TCTTACAGAA CAAGGCTATA CTGAGAAATT CAAGGAACAC CTTAATGAAT
66951 CAGGGTTATT TCATTGTGAG AGAGTATAGA AATGGTTGAT AGTATGCTGT
67001 AAGTAATTTT ACTTTACGTG TAAGTACTTT TCTCTGGCTT TCCAGAACAT
67051 GCTGTTGGAG TAAGAGAGGA ATGCCTTATT GTGGACCGAG GGGATAGATT
67101 TGGATACAGC CTTCTTTGAG GAAGGTAGAG AGGTCAACTA TTACATATCC
67151 AGCAGTAACT TCCCTTTCAA AGACTAAGTG TTCTCATTCA ATCATTTGAT
67201 ACTTTTTGTG CATCTACTAC AATTTGAAGA ATAGAAGAGG GAAATGCATG
67251 TGTAAGGCAT GGTGGCAACA TTTAAGAAAC CCAGATTTGG GGATACAAGG
67301 TGTGTGTGTG CACATGCATG TGTGTGCAAT TTAAATGCAC AGGGTAAAGA
67351 ACTTTGAGTC AGTGACAATG AGTGGCAATG GTGGTAATTA AGTGCTGGGG
67401 AAGATCAAGG GAGAGAGAGG GTGCTGAGAG CCAATAGGGT GGAACATGTT
67451 TGAAAGAGCT GGGTTCTGAA GTATTCCTCC ATAGAAGGGC ATTTTAAATA
67501 GCTTTTTTGC GTCCTTATTC TGTTAAGCAT TATTAAATTG TTCTCCCATC
67551 TTTAAAAGGT TCCCTAGCTT AGGTGCACTG GCAAGAAATT ATAAAAGCAG
67601 CATGACCAGG CGTGGTGGCT CATGCCTGTA ATCGCAGCAC TTTGGGAGGC
67651 CAAGGCGGGA GGATTTCTTG AGTTCAGGAG TTTGAGACCA GCCTGGGCAA
67701 CATGGTGAAA CCCTGCCTCT ACAAAAAATA CAAAAATTAG CTGGGAGCAG
67751 TGGCACGTGC CTCTAGCCCC AGCTACTCAG GAGGCTAAGG TGGGAGGTTG
67801 GCTTGAGCCT GGGAGGTGGA GGTTGCAGTA AACTGAGATT GTGCCACTAC
67851 ACACAGCCTG GGCAACAGAG CGAGACCCTG TCTCAAAAAA AAAAAATTAA
67901 ATTAAAACAA TAAAAGCAGT GTGGTTGCTT ATAAGGAGTG ATGGAGTGGA
67951 CAGAGGAGGT TTTTGGGCTA GTCAGGTAAG GCTGGGGTAT GAATCTAGAA

FIGURE 3-20

68001 GTTTCCATTT AAATAGCAGG GAGCCCTTAG GCAAATCACT TAACTTCTGA
68051 GCTTTGCCTG TTTCACCTGA GGTTGCTTCA AAGATTGAAT GAAACCGTAT
68101 ATATAAAGTG CTCCTAAACA TCATCTGCCA TGTGGCAGGT TCTCAAGAAA
68151 TGTTAGTTTC CCTTTCTCCT TACAAAGATA AGATGCTTGC TTTGAGTATA
68201 TTTTTAGGCT TCCTGATCAT TATTGGTTAT GATTTTAAAT CTTGGTCCAG
68251 CCACCACCCA TATGGTTTCT GCAGTTAACA AAAGAGGCAA AGGTTCACTA
68301 CTGGGGGAAA AGAGGTGTAC AGAAATGGGT GTAGAGAAAG TAGATGTTTG
68351 AAGGGGATCT AATTAGGAAA GTATTTTTCC TGGTGGTCCA GAATTTAAAA
68401 TTATAGAGTC TTATGAGAAA ATGATAATAT TCAGGTTAAG AACAGTTTAT
68451 TATTTCCTCT CTATATTGGA GAATATTTTC ATTATCTTAT ATAAGAACCT
68501 TGTAAAAATT TTTCTTCTTA ACTAGCTTCC CACTATAGGC CATTAATCCT
68551 GTTATTATTT CAAAGATTAA ACAACTATAG TAATAACAGG ATTATATGTG
68601 CATTAATTTA TATTACTTCA TTCCGCAAAG GATTTGAGGT AGTTTTTAGA
68651 AGCATAAAAC ACTGCACAAA TAAAATAGTA GGATAGGAGT AGAAAATAAA
68701 ATTTCAGCAA CCATGAAAAA TATGACATAG TATATGTTGT TAAGACTGGG
68751 GGAATGTAAA CACATCACCA GTCAGGGCCT ATATAGTTGT TACAGTCCCA
68801 TAGCAAATTT GACTCTAAGC TTCTGGCAAA CAACATGAAA AGGGAAGCAT
68851 GATGGATGTG GTTAAATAAG ACACAATTTA CTTGTTACTT TACTCAAGGA
68901 AGCAAGTATT TTTTGCCCCT TTGTTTCTTA TAGGAGATGC TGGGTAATGA
68951 AGTAATGTTG GCATTGGCCT TATAGTGGAG GTAATAATGG ATTTTATCAG
69001 GCAGTTTCTT TAAGCATCTC TTGATGAAAG ATGAGGCTAT GACATCAAGA
69051 GACAATTCTG AGGCCGGGAG CAGTGGCTCA CACCTATAAT CCCAGCACTT
69101 TGGGAAGCTG AGGCGGGCAG ATCACTTGAG GTCAGGAGTT CGAGACCAGC
69151 CTGGCCAACA TGGAAACCTC GTCTCTACTG AAAATACAAA AATCAGAAAC
69201 CCTGTCTTTA CTAAAAATAC AAAAATTAGC TGGGCGTGGT GGCAGTGCCT
69251 GCAATCTCAG CTACTTGGGA GGCTGAGGCA GGAGAATTGC TTGAACCCAG
69301 TAGGTGGAGG TTGCAGTGAA CTGAAATCAC ACCACTGCAC TCCAGCTTGG
69351 GTGACAGAAC AAGACTCCAT CTCAAAAAAA AAAAAAAAAA AAAAAGACAA
69401 TTCTGAAAGT GGTGAGTCGT TCTACCAGGG GCCAGGATGA TCTCATCTGG
69451 GTTATGGATT CTAGGTCTGG CCTCATCTAA GGTGACACAC AGAGGGTACA
69501 CCGCACAATC TTCCTATCTT CTTTAAATAT CACTGCTTTC AACAGGACTT
69551 TTTTTTTTTT TTTTTGAAAC AGGATCTTGC CCTGCCCTGT CACCCAGGTT
69601 GGAGGGCAGT GGCATGATCA TGGCTCACTG CAGCCTTGAC CTCCCAGCTC
69651 AGCCTTCTGA GTAGCTGGGA CAACAGGTGC ATGCCACCGT GACTGGCTAA
69701 TTTTAAAAAG TTGTTTCGTT TTTTTTTTTC TGTAGAGACA GGGTCTCCCT
69751 AGATTGTCCA GGCTGGTCTC AAACTCCTGG CCTCAAGCAG CTCTCCAGTC
69801 TTTGCCTCTC AAAGTGTTGG GATTACAGAC GTGAGCCACC ACGCCTGGCC
69851 CAAAAAACAT AATAATGTGG TTATTCGAAG AAGTGATTTC CTCTCAAAAC
69901 ATAAATTCAT TTCTTCTTTT ACTCTTGTAA CTTTCTGAGG TGAAAAGTAG
69951 GAAGTTCCCA GATTTTTCAT TGCTGTAAAG AAAGATTATA GACAAGCAAG
70001 GAAGGAGTTA GAATAACCTG TGTGATAATG AATTAGAAGA GTTGGAGGTG
70051 TATGTAAGTA TGCTCAGCAT GAATTTATGT TTAGCTTAAT GTAGATACAG
70101 ATGGTTACAT GTGGAAAGTA TTTATAGCTA TGCATAGATA GGTTGGTATA
70151 TGTACATGTA TTTTCTACCT CCTTTGGCTA AGAGGGCGCA GAAGCCATGA
70201 CATCCCTGTT GCAACAAAAA CACCTAGCAC CATTTATCTT GGTTTGTAAT
70251 ACTATTCTCC AGTAAAAACA ACCAGGGCTC CTTGCAGAAA GGGCTGATGA
70301 TAATATATAA GATTAGCCTG GAGCATCTTA TATATCAGAA AGCAAGGGAG
70351 TACTCAAAAA CTAAAAACAA TAAACTGCTC CCCAATAATG GGAGTATGTC
70401 AAAGGGTCAC AGGAGCCATG TGAAAGAGTT TGCAATAGCC AACAAAATGA
70451 AGAAGTATTT GAATTTAAAT CAGAGTATAA AATAAATATC TACGAGTCCA
70501 TAATGATATA AACAAGTGAT TGAATAAATA AATAAATGTG GGAGACTAGA
70551 CAAATCCCCC ATGCAGAAGA ATTCCAAGTA ATTTATGTAG GTAAATACTT
70601 CACTGTCAAA GAGGCAGAGC ATAATTAATT CCCCACTCCT GTAAGTGTGG
70651 GCTGCTTAGT GACTTCCTTC CAGGAGTGCA GTATAAATAG GGAAAAGAAG
70701 ACAACTTCAC AGTGGAGGAA TCTGGCAAAC TCTGTCCCAC CCAGATGATC
70751 AAGGTTTACA TCAAAAGCAC TAAGCCATGT TGATACTGCA CTGTGTTTCC
70801 TTGATATAAT GGGATGAAAT GGCACTTTGT GTAGTCGTGC TTCCAAAAAA
70851 CCTCATAGCC CTAGACTAAT CATGGGAAGA ACACCAGACA AATCCTAATT
70901 GAGGGACATG CTACTTAATT CCTGAAAAGT ACTCCTCAAT GCTGTCAAGG
70951 TCATCTGAAA CAAGAAAAGT CTGGCAAACT GTCACAGTCA AGAGGAGACT
71001 AAGGAGATAT GACGACTAAA TATAATGTGG TATCTTGTAT GGGATCTTGG
71051 AATAGAAAAA GGATATTAGT TAAAACTGAG GAAATCTGAG TAAAATATAG
71101 ATGTTTGTTA CTAATAATGT ATCAATAGTA GTTCATTAAT TGTGACAAAT
71151 GTAACATACT ATCGTAAGAT GTTAATAATA GAGGAAACTG GATGTGAGGT
71201 ATATGGGCAC TCCCTGTACT GTTCACAATT GTTATGTAAA TCTGAAACTA
71251 TTCTAAAATG AAAGTGTATT TTATTTTATT TTATTTTGAG ACGGAGTCTT
71301 GTTCTGTTGC CCAGGCTGTA GTGCAGTGGT GAAATCTCGG CTCACTGCAA
71351 CCTCCACCTC CTGGTCTCAA GCGATTCTCC TGCCTCAGCC TCCTAAGTAG

FIGURE 3-21

71401 CTGGGATTAC AGGCACGCGC CACCACACCC AGCTAATTTT TATATTTTTA
71451 GTAGAGATGG GGTTTCACCA TGTTGGCCAA GCTGGTCTTG AACTCCTGAC
71501 CTCAGGTGAT CCACCTGCCT CGGCCTCCCA AAGTGCTGAG ATTACAGGCG
71551 TGAGCCACCG CGCCCAGACA AAAGTTTATT TTTTAAAAGG TAGGAAAGTT
71601 TCACATTTTG ATCGTACTAT TGAGATAAAA CTTGGTTGTT GTTGTTGTTT
71651 TTGAGATGGA GTCTTGTACT GTCACCTGGA CTGGAGTGCA ATGGCATGAT
71701 CTTGGCTCAC TGCTACCTCC ACTTCCCGGG TTCAAGCGAT TCTCCTGCCT
71751 CAGCCTCCTG AGTAGCTGGG ACTACAGGCA CCTGCCACTA CGCCCAGCTA
71801 ATTTTTTGTA TTTTTAGTAG AGATGGGTTT CATTATGTTG GCCAGACTGG
71851 TCTCTAACTC CTGACCTCGT GATCCGCCTA CCTTGGCCTC CCGAAGTGCT
71901 GGGATTACAG GCATGAGCCA CCGTGCCTGG CCAAACTTGG TGTTTTCATC
71951 TGAGAATGAG GTTTCCAAGA ATTGAAGATA CAAGTTAGCT AAATAGTATC
72001 AGTGAAACCA CAGAGTATAA CTTCAGACCA CTTGTGGTTT TAAGGCAGAT
72051 CTATGCCACA GAGAAGTATT TGAATTTAAA TCAAAGTATA AAATAAATAT
72101 CTATGAGTCC ATAATGATAT AAACAAGTGA TTGAATAAAT ACATAAATGT
72151 GGGAGACTAG ACAAATCTCC CATGCAGAAG AATTCCAAAT AGTTTATGTA
72201 GGTAAATATT TCAGAGTATC AGTACTATGA TCCTTCTTAC AATGCTATAT
72251 CCGACTAAAA AAAAAAAAAA GAAAAAAAGA AAAAGAACTG TGATCCTAGT
72301 TGCGAATGTT GACTTATAGC CAGAAAAGAT GACTTTCTTA TTTGAAAGCA
72351 TCATGAAATG CTATCAGCAT GCAGAAATGA GAGGGCATAA CAGAGAGCTA
72401 TTCTATCGTA TTTTTATTTA ATGTTTTAAA AGAGCTATTA ATAATATAAT
72451 GTGAAGAGAA GTTTTAATTT CAGAGGAATG GGATGTTTGG TTTTAAGATT
72501 ATTTTTGCAA AATGTAGTCT GTCATTTTTA AAGTGAACTT TATAATGAAA
72551 AGAGTTTGAA ATTATTTCCC TATTATTAAT ATTTATTTAT TTTTATTACT
72601 TTTCTGCCAT GACAATTACA GTGCACATTT TCCCTATTTT TAGCTTTTTT
72651 CTGGTAAAGT GATTTTACCT CCTGTAGATC GTCAGTCCCC AGACTTTCTG
72701 GCACTAGGGA CTGGTTTCAT GGAGGACAAT TTTTTGACAA ATTGGAGGGG
72751 GAATGATTTC AGGATAAAAC TGTTCTATCT CAGATCATCA GGCATTAGAT
72801 TCTTACAAGG AATGTGCAAC CTAGATCCCC TGCATGCACA GTTCACAATA
72851 GGGTTTGTGC TCCTATGAGA ATCTAATGCT GTGCTGATCC GACAGGAGGT
72901 GGAGCTCAGG TGGTAATGCT CGCTCACCTA CCACTCACCT CCTGCTGTGT
72951 GGCCTGGTTC CTAACAGGCT CGGGGGCTGG GGACCCCTGC TGTAAATAAC
73001 CTTTCGAAGA TCTGAAAATT AACTTTAGTA TTTTTTTGTA TTTACTCGAT
73051 ATTTTAAACA AACAAAAATC TAGAGAATGA CATAACAAAT GTATTTTCTG
73101 CGTATCCACT ACTCTTTACT CAGATCCTAA CATTCTCAGT TTTGAAAAGA
73151 TGTAGGCTGG TCACAGTGGC TCATGCCTGT AATGTCAACA CTTTAGGAGG
73201 CCAAGGAGGA CAGATCACGT GAATCCAGGA GTTAGAGACC AGCCTGAGCT
73251 ATAAGGCAAA ACCCTGTCTC TACAAAAATT AGCCAGGTGT GGTGGCCTGC
73301 ACCTGTAGTT CTAGCTACTC GGGAGCCTGA GGTGGGAGGG TCAATTGACC
73351 AGAGGCTATA GTGAGCTGTG ATCCTGCCAC TACACTCCAG TCTAGGTGGC
73401 AAAGCAAGAC CCTGACTCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAG
73451 ATGTAAACAT TACAGGCCCA GCTGTGGTCC CGTGTACACT TCTCCATTCC
73501 TAATCCCCTT CATGCACACT TTCCCACAGG TAATCACTAT CCCACATTTG
73551 CTTTAATTAT TCCCATCATG TTTTATGCTA TGACTACAAA TGTGTGAATC
73601 TATATTCCAC TGTATGTTAA TACCAACATT TTTTCATTTA TTTTACCGCT
73651 GATGGGCAAT TTAGGTTATT TCCAGATTTT TACTATTACA GATGCTTCTG
73701 TGAACATCCT TACTCATTTC TCCTTGTGCA CATGTGCAAA ATTTCTCTGG
73751 AGTACATACC TGGGAGTGGA ATTACTAGAA TGTGGTGCAT GTACATGTTG
73801 TTGTATTTAT CAGTAGTTCA CTCTTATTTA TTGCTGTGTA TTCCATTACA
73851 TAGATATACC ATAATTTGTT TATCTAATCA CTTATTTATA GACATTTGGG
73901 TTCTTTCTGT TTTTTTGACT ATTAGAAATA AAGCAGCTAT AAACATTTGT
73951 ATATAAGTTT TTGTATAGGC ATATGCTTTG GTTTCTCTTA GGCTCAGGGG
74001 CTGGGAACCC CCTAGGAATG GAGGAATAGC TGGGTCATAT GGTAGGTTTA
74051 TGTTTAACTT TTTAAGAAAT TCCTGAACTG CTTTCCAAAG TGTTTGTACT
74101 ATTTTATATT CCCAGTAGCA GTATGTGAAA GTTCTTGTTG CTCTACATCC
74151 TCACCAACAC TTGGTATGAT TAGTCATTTT TAGATATCCT GTGATGTGTG
74201 TGGTTTTTCA TTGTTTTAAT TGACAGTTCC CTGATGACTA ATGATATTGG
74251 GTGTCATAGG CATATTTTCC ATTTATATAT CTTTTTTGTA AAGTTCAAAT
74301 TTTTTGCCCA TTTTTAAAAT TGAGTTGCTT GGTTTTTATT GAGTTCTGAG
74351 AATTTTGTAT GTAATTTATA TACAAGTTCT TACTAGATAT GTAATTTGCA
74401 AATCTTCTCT CTCAGTTTGT GGCTTGTCTT TTTATTCTCT TAGCAGTGTC
74451 TTTCAAAGAG AAGTTCTTAG TTTTGATGAA GTTTAGTTTA TCAACATTTT
74501 CTTGTACTGA TTTCTTGTAT GGGCATTTAG GTTTTTCATG TTATATCTAA
74551 GAAATCTTTG CCTAAACCAA GATCACAAAG ATTTTCTCCG GTATTTTCTT
74601 CTTTTGTTTG AGACAGAGTC TCACTCTGTT GCCCAGGCTG GAGTGCAGTA
74651 GTGCAGTCTC AGCTCACTGC AACCTCCACC TCCTGGGTTC AAGCGATTCT
74701 TGTGCTTCAG CCTCCTGAAT AGCTGGTATT ACAGGCATGT GCCACCATGC
74751 TCAGCTAATT TTTGTATTTT TTTAAGTAGA GACAGGGTTT CACCATGTTG

FIGURE 3-22

74801 GGAAGGCTGG TCTTCAATTC CTAACCTCAG GTGATCTGCC CAACTCAGCC
74851 TCCCAAAGTG CTGGTGACAC AGGATTTTGC TCAGCTACTT TGCCAACCAG
74901 GGACTCACTC GGCCATGGGC CAAGGCACCC CGCTCACTTG GTCCACCTGT
74951 GCTATAGCTT CTACTGACGT TCAGCGGTTC CCGAGCTCTT GTCATGCATC
75001 TAAAAAGAAG GAGGATATGC TGACAATTTG AAGTGTAAGG ATGGGTGGAG
75051 AAGAATTTTA CTGAGTTATG GAACAACTCT CAGCATTAAG GGGACACGGG
75101 GTGCTCCCTC ACCCCCACAG TCAGGTGGTT TTTCTCTCTC TCTCTCTGTG
75151 TCTGGGTCTG GGGCTTTTTA TGGACTCAGA ATGGGGAGTG TGTACAGATT
75201 GGTTTGTGAG TATGCAATAA AAGTTAAAGT GAAGACAATA TTCAAAGGTG
75251 GGCGCGATAG TGTAGAAAAC CAATTAGGAA AGGGTAGGTA TGTAGCCTGG
75301 CATGGTGGCT CGTGCCTGTA ATCCCAGCAC TTTGGGATGC CAAGGCAGGT
75351 GGATCACCTG AGATCAGGAG TTTGAAACAA GCCTGACCAA CATGGTGAAA
75401 CCCCGTCTCT ACTAAAAATA CAAAAATTAG CAGGGTGTGG TGGCACACAC
75451 CTGTAGTCCC AACTACTTGG TAGACTGAGG CAGAAGAATT GCTGGAACCT
75501 GGGAGGCAGA GGTTGCAGTG AGCCAAGATT GCACCATGGC ACTTCTGCCT
75551 GGCTGACATA GCGAGACAGT GTCTCAAAAA AAAAAAAAAA AAAAAAAAAA
75601 AGAAAGGGTA GGTATGTGTA AAATAGGTGG AGGGTGGGGA TCAATCAGAG
75651 GAAAGCATGC CAAATGGGAA GACAGGTTCT CAATCCAGTC CATGGATTTG
75701 CCTGGGACTT GTAGCTAGGC TTTAAAGTGT CTTCAGCTTG AAGGTCAGGT
75751 TTCACCAGGG ACCCGTTCCT ACACCCCTAT CTGCCTAGGC ATTTGTCTGC
75801 CTCTTGCTGC TATCAGTTCT CCCCTCTGAA GAGGTACATC TAACTGCCAT
75851 TAGAATAGGG ATGATGACTG ATACTAACTG CTTCCTGCTG ACAGGGGTGC
75901 GGGGTGCTGT TTTGGGAAAA TGGCAGTCAG ATCTCCCTCA GAGGCCTATC
75951 TAAGGGTCCC CAGTAAAAGG GAGCCATCGT CTGAGGCTCC AGTTTCATGA
76001 CTGGAGTTTA ATGGCCTGAA AATGAGAAGA CAACCAGATT ATTAGAAGGC
76051 ATGTATCAAA ACAAAATAAG GGGGTAAGGA CAGCTCAAAA ATCCTGAGGC
76101 TGCCAACATG CCCAGATAAC AGGTGGCTAT AGTTATGCCT GCTAAGATTT
76151 GGGTGAATGA GGCTCGGCTT TGGTCAGCTT CTTTGGTCTT ATTTTCCCAA
76201 ACAAAGAAAC CTCTGGGTTA CGGGCACCCT GTTTACTCCT ATCACCTGGC
76251 AGGATTTGCA GGATAATTGC CCAGAACTAG AATATTGATC CAGATTTTTA
76301 CATCACCCAT CCCTTTTGTT TCTTCTGAGC TGCAGCTGAT GATCACTGGT
76351 TGGTTCACAG AAATAAGCAG GGTTAGTCTA AAATGCAGAC AAAAACTTAA
76401 AAACAACTAA TGAGACTAGA ATTTAATGAA AAGTGTATGA TAAATTTTGA
76451 AACATAATTT TTCTCTCTCC AGTCCTCATT TTTGTTAAAA ACAAATCATG
76501 ATAGGACTGA GTCATTTGCA GAATAAACTT TAGTCTTATA TTTGGCCTGG
76551 TTATTTGCAT AAAGCACAGC AAGAATAATT ATTTTTCACA CAGGCTTTTA
76601 AAATTGGCTT TGATGGAACT CTGTTCCACA AGGAATTTCA GATAAGACCT
76651 TTTAAAGCTG AGCCCAGCCA TGGGTTTGTA TCCTCAAATA CCTATGAGTT
76701 GGGTAAATTC CTCTCTTCTT GAGGTCCCAA GATAACATGG GGTTCCTGGG
76751 CCTATTAGAA AGTGACATTC TTTATTCACC ACAGATTAGG AACTCTGTAC
76801 AGGGACTGTG TAGAAGACAA AGTATGAGGC CAGTTTTCCC AAGGGGCTTT
76851 TATTGGTTCT GCAAGTCAAA CTTGATTTCT TAAAGGTAAG CACACCCTTC
76901 CAGTCAAAGC CTTGGTAAAA CAACCAGTTT CTCCAGTTGT GTCCTGTTGC
76951 AAAAGAAAAT GGATTCTTAC TGCACTGATG CAAATAACTG TATTGCTGCA
77001 AATTAAGAAT ACTCACAAAT AGTTTCCAAG TTCTGAGGAA ACCAGGCAAA
77051 AAGAAATAAA TGTGCTCCAA ATTTTGTTCA CTGGAGTATA CCTTACTCAA
77101 TTGTTAAAAG CTATAGATAG CTCAACATGA AAGTTTCCTT GACTCTGAAA
77151 AACAAAACAA GGATCAGCAA TGTTTTAAGC AAAAAAAGAT TACTTCAGCT
77201 TTCTATTAGT TCAGTACATT CTATTAACTC TTCTTCTGCT TGATATTCAT
77251 GAACATTTCA GCTCTTCATG AGTCCTGTAC ATTTTCCCTC TATTCCAATG
77301 TCACAATCTC CAAAGTTATC AGAAACCTGC ATTTGAGAGC ACCTGTCAAA
77351 GTCCCATAGC TGATTATAAA CCATCTTTTG AAAAGGATCA AAATAAGACA
77401 ATTGTCTGTG AATGACAAAA TGTCTTTGGG TAATAACAGT CAAAGCCATG
77451 ATTGACAAAG AAATTTGGTT ATTTCTGAGC TTTACAATAA CAACATAATA
77501 ATTTTTTTTT TTTTTTTTTT TTTTGAGACG GAGTCTCGCT CTGTCGCCCA
77551 GGCTGGAGTG CAGTGGCGGG ATCTCGGCTC ACTGCAAGCT CCGCCTCCCG
77601 GGTTCACACC ATTCTCCTGC CTCAGCCTCC CAAGTAGCTG GGACTACAGG
77651 CGCCCGCCAC TACGCCCGGC TAATTTTTTG TATTTTTAGT AGAGACGGGG
77701 TTTCACCGTT TTAGCCGGGA TGGTCTCGAT CTCCTGACCT CGTGATCCGC
77751 CCGCCTCGGC CTCCCAAAGT GCTGGGATTA CAGGCGTGAG CCACCGCGCC
77801 CGGCAACATA ATAATTTTAA TTACGATTGA TAGCATATAC TCAGACATTA
77851 GAATTTTAGA AACCTCATAG AATTTTGGAA CATATGTATT TTTCATTAAA
77901 ATATAACCTG AAGAAGATTA AACATTATTT TTATTTTGGC AATCCCACAT
77951 AACTAAACAT GTCAGTTAAT CCTGTTTACC TCTCTTTTGG ATGCTCCAGG
78001 AGCCCTCTGT AGTATTCAAA AGTAAGGGGT CAGAAAAGAC AACCTTGAAA
78051 CTGAAGTTTG ATTTTGGGAA GCCTGTTAAG TACATTAGAG GTTTAAAACA
78101 CTTTATATTA TGAAATACAA TTCCAGATTA CCATAAGTCA TTTATTTAGC
78151 CAAAATGATG ACTCAAAAAT TTTTAAAAAG GTAAAAACCT TTACTCATTA

FIGURE 3-23

78201 AGAGTGAAGA CAGCTTTCCA AACAAACAAT CCATCTCTGG TCTCTCCCAC
78251 ACCCTTTATT TTTTGATGAA ATCTTTAGAT AATTCTGTCC AATCTTAACC
78301 AGTTTGACCA TGAGGTGAGA TTCTTATAAA CCTTTACAAA TTTTTGTTAA
78351 AGAGTAGATC AGTGCCTTAA GAAAACCTTG TTCTTTTATT CTAATGTTCA
78401 ATTTACAGAA AAACCATGTA ATACCCTTTT GAATTTAGTC AATATGTTCA
78451 CACACAGAAT TTCTTTTGCA AGATTAATTT TTACAAACCT TCCACAACCT
78501 TCAACTTTAT CTTATCCAAC TTAAAACAAT TATTTAATCC TCTAAACTAG
78551 GCAAAAATTT AAATTCCCAT GCCTTCTTAT AATCCTTTAC TAAAAACACA
78601 TTTACTTTCC TTACACACCT TGATGTAAAT CTGTTTTCAG TAGTCTCAAT
78651 TACATGGTAT AATGGTAAAC CTTAGCAATT TTTAATTTTA ATGTAAAGCC
78701 TGGTAAGTTA TTTTAATTAT GTGCTACCAA TTATACCTTA AGTTGTAGCG
78751 ACTCTGGTGT GCTATTGGTA ATATGGCCTT ACAAAACTGA AAAGCAAGCA
78801 AGGTGAACAA TTTTCAAAAG CCCAAGAAGC AGGCTGGGTG TGGTGGCTCA
78851 CACCTGTAAT CCCAGCACTT TGGGAGACCC AGCCAGGTGG ATCACCTGAG
78901 GTCAGGAGTT CAAGACCAGC CTGGCTAACA TGGTGAAAAT CCGTCTCTAC
78951 TAAAAAAAAA ATAAAAAAAT TAGTTGGTCA TTATGGTGTA TGCCTGTAGT
79001 CCCAGCTACT CAGGAGGGTG AGGCAGGAGA ATCACTTGAA CCTGGGAGGC
79051 AGAGGCAACC AAGATCATGC CACTGCACTC CATCCTGGAC AACAGAGTGA
79101 TACTGTTGAA AAAACAAAAA AGAGAAAAAA GCCAAAGAAG CAGTTTATAA
79151 CTTTAAAGCA TTTAGTAAAC CTAGTATCTG ACCTGCATAA TTTAGACCAC
79201 ATGTTTACAT TTTGAAGACA TTTGTATTTT ACCAATAATC TCTAAAACTT
79251 TTTATTTTTC AAATATTAAA ATCATTTGAA CTAAAAGGTA TTATAGCTTT
79301 TATTTTTCCT TCAGAAAATA TTTGATCTAA GTGCTTATTT TTCTCTAAGC
79351 CAATTAATTA GAGCTCTTTT TTATACAAAC ATCACACATA TTGCACATAT
79401 ATAACTACAC AGACAGAGGA TCCAGTAGTT GTAAGATTTT TCATTGTCCA
79451 ATCTCCTAAT TAGATTACTG ACCTCAGGAT GGAGCCCTTC AAGAGCAGGG
79501 CTAGGAAAGC ATGCAGTTTC TAGGGCCTAA TAAATAGTTA TAGCTGGAAG
79551 ACAAAAACAG ATTTTGAGAG GGATTTATCT GCTTTTAATT CTTTGGGTTT
79601 CATGAGGAAA ACAGAGGTTT TTTTCTAAAA TGGGGTCAGT GGTGCCTCTT
79651 CCATTTTTTC CAGGGAGTCC CAGGCCATCA GAAGTTATCT TAGGGCCTCT
79701 CATGCGTGCA TTAAGAGAGG CAAGACAAAA TGGAGAAAAG TAATTCAGTT
79751 GACTGAAAAA GAAAATCTTT TTCCAGTGAA ACAAGATGCA AGAAGAGGAA
79801 AACATAGAGG CCTTTTAAAT ATGCCTATAG CTTGGATATC CACTTTTAAT
79851 TAAGCTGACT TTTTACCATA GTGCTCTTAT TTTAAAAAAT CCTTTTAAAT
79901 CCCTTGTTAC CCAACTTTAG CCTCACCAAG TGGCCAATAT TTCTGGCTTT
79951 TGAACTTTAC CAAAAGTAAC CTCACAGGTG AAACCAACAA GCCTTAAGTA
80001 AGGTTGTGAC TTCACTGCCT GTGTACAAGG TATTTTCAAA GAGATGGTAA
80051 GCAGTTTTTA CAAAATCTAG AATCTTTAAA GATAGCTCAG AGAAAGGAAG
80101 ATTTAAGAAA GGAAGCTAGA AGTTGTTCAT GGAGGGGAAG AGAATCAGCA
80151 AATGATAAAA GTCACACAGA TATTAGCCAG AAAGTACTCA TTCCCTAAGC
80201 CAGGATTGAA CCTGGGCTTC CATTGTAAAA TGGCAGAGAC CAAAAGAAAG
80251 TCCTTCCACG TGGTTACAAG GTCAAGCTCC CAAGGACATA AAACAAGATG
80301 GAGACCTTAT TCAGTTTTTT TTCCTTCAGA GACCTGTAGC AAAGTCTGTA
80351 ACTGACCAGT TTGCATGGCT GGCTTGAACA GTGGGCTTAT GGGGTCCTAG
80401 GTCTGTGTTC TATCCTGTCT TACTCCTTAT GACAGAACTG TACAGAAAGA
80451 CACACAAAGC ATAGCAGATT GGCTACAGCT TAAGATTAGC CTCACAAATC
80501 CTTTTTTCCA TTAATCAAAA CTTTACAGAA GAATAAACAG TGATTTTTAT
80551 CCTTCCTTTT ACTGGTTTGC ACAGGGAGAA AGAGGCCAGA AGTCTGACTG
80601 GTAAGAACTT TTACTCTTTT ACTGGCATGT CAGGCTTCTG GGTTCCCTTC
80651 CCAGTTCAAT TTTAAGCCAA GCAGTTTAAG GTTTGGGGAT ACTAAACTTT
80701 TCACAGATAT TTTCCAGTAT GTTTAAATAG TTTTGTTTAG CCCAGATTTA
80751 ATAATAGTTA TCTGTAGTAG GGTTTGCATG ACCACTTTAT TCCAATGCCA
80801 CAAACAGAAG TTGATTAGCA AATTATATAT ATATTTGCAT ACGTTTATTT
80851 TACATAATTT AAGGCTATTA GACCAATATT GGTTATTACA GAGCATAGAT
80901 ACCTTGAAAG AGGCCAAGAT GGAGTGGAAA TGCAGGTTGG AGCTACCAAA
80951 GGGCAAAAAA CTATTCATCA ACCTTTAACT ACAATAAATC TTCCCTTTGG
81001 GTATTTTCAT TTGTAAACAC TTTGCCCTTG CAGTCATTTA CATCGGCATA
81051 ATTACAACAC CTTTGCTTTA TTTGGCATGC AGGAGAATCT TCTTAAATAT
81101 TAAACCTAAC CACTTTGAGT GATTTGCATC CTGCTTTTGT CATACATTTA
81151 GAGCCTTGGG ATTTTTTCTC AGAATGGAGT AGGAAACAAA TAGGGTCTGG
81201 ATAGGGAAAT TGAAAAGCTT CCTGGTATTT TTCTGTTGAA AGATTTCTTA
81251 ACATGGCTTC TTGGATGTGT CTCTCTGATG TCAAACATAC ACACATATTC
81301 AAATAAGAGT TATACAAGCA CATCTTGCAC ATTTTTGGCA TCTATGTCTC
81351 AAGACACAGG ACATTCTATC TGGTGTTCTG ATCGAACCAC TTTTGCATGT
81401 TACTAGACTG AAAATTATTG GAAGGTAGAG AAATCTCTTA TTTGTTTTTA
81451 TATTCCCAAC AGCCTAGGAT AGAGCCTAGA ACATTAAAGA ACACTAAAAT
81501 TTTAATAGTG TAACTGAAAA GCAGGTTAGT TGGTCACTGC ATGTAGAGTC
81551 CAATTAACAA GAGCAAGTTC TGATACAAAG AAGTGTATTT ATTTCAAAAC

FIGURE 3-24

81601 TAGCTTAGGG GAAGAGGCAC AAAGCATCCT GCCTTTAAAT GTGCCACTTC
81651 ACCTTTGGAG CAAAAAGTGG GCATTTTTAT AAGGTAGGGG AGGAAATGAG
81701 CAAGGGCAAG TGTCCCTCTG CTACTGGGCA AGTATCTGAG CTGGCACCTT
81751 CTTGGGCAGA AGTAAGTTGT AAAAGTGGCC AAGTGGGTAT GCTTTCAACA
81801 TGCCCTCCTA GTGGGCATGA GTTCTGAGAT GACCCTGTGG AGAGTTCTGT
81851 GGGGGCATGC TTTGGTCTGC AAATAGACTG TTAACTTTCG AGGAGAGATC
81901 CTTGGGGGGA AAATATATAT TAGGAAGTCC TCTGTGGGTG TTTTGTAGAA
81951 GGACCTAGAG GGACTAGGGC TCGATTGTTA TTTATTTATT TATTTATTTA
82001 TTGTGTGTGT GTGTATGTGA GAGAGAGAGA GAGAAAGAGA GAGAGACGAG
82051 GTCTTGCTCC GTTGCCCAGG CGGGAGTGCA GTGGCATGAT CATAGCTTAC
82101 TGCAGCCTCA AACTCCTGGG CTCCAGGGAG CCTCCTGCCT CAGCCTGCCA
82151 AGAAGCTGGG GGTACTGTTG TGTGCTACCA TGCCCAGCTA GTTTTTAAAG
82201 TATTTTTTTT TGTACAGATG GGGTCTTGCT TTGTTGCCCA GGCTGGGCTT
82251 GAACTACTGG CTTCAAGTAA TCCTCCTACC TCAGCCTCCC AAAGTGCTGG
82301 GATTATACAT ATGAGCCACT GTTCCTGGTC TAGTTTGCAC TTTTTTTTTT
82351 TTTGAAACAG AGTCTTACTC TCTTGCCCAG GCTGGAATGC AGTGGCACAG
82401 TCTCAGCTCA CTGGAAACTC CACCTCCCAG ATTCAAGCGA TTCTCATTAC
82451 TCAGCCTCCC GAATAGCTGG GATTACAGGC ACCCGCCACC ACACCCAGCT
82501 AATTTTTGTA TTTTTAGTAG AGATGGGGTT TCGCCCTGTT GGCCAGGCTG
82551 GTCTCGAACT CCTGACCTCA GGTGATCCAC CCACCTCGGG CTCCCAAGGT
82601 GCTGGGATTA CAGGTGTGAG CCAACACGCC TGGCCGTTTG CATTTTTAAG
82651 ATAAAAATTT TACCATGCTG GATATATTGT AGTAGCTATG TACTTCAGTT
82701 TCTCAATTGT TAAACGAACT TAATAGAAGT ACTACCTTAT AGAATTTTGG
82751 GGATTAAATG AAATAGTCTT CTGAGCACAA CAAATATATT ATCAGCACAA
82801 CAGCTAGCTC TATTGAGTCT TTATTATTAT AATAGCAGTA ATAGTCAGAC
82851 TTGGAAGGGG TGAAAGAGAA CAACAGTCCA TTTTATTTTT GTGGCATATA
82901 TATCATAGGT CGTAAGACCT TGGATTGTTT AGATGCCATG TTAAAACTTG
82951 ACAAAACTAG AAATGTTGTG AGTGTGCAAT AGCAGGTGAT AACTAATCCA
83001 ATCATTAATT TATTCTTGAA TTTGATCAGA AGACAGACCT AACTTCATCT
83051 ATTGCCAATT ACTATTGTAA CAAAATCTAT TGGAATTTCA GTTTAGGCAC
83101 TGCAGTACAA CAGTGTGAAT TTCAAAAGTG AGATATTTTA TGTGGCTTTT
83151 TAAAGTAGGT TTTCAAACCA GTTAAAGGCT CTAAAACCCA TTAAGAAGGA
83201 TACTATGGGT CAGGAATAGC ATTTTATGAG GACTCTTGAG AAAACGGTGC
83251 TGTGTCCCCT CCACCATGGT ACAAGTGGAG GTTACTCTCA CTTGTTCTCT
83301 TACTTGGTAT CTCTCCTTGT TCTCTCACTT GGTATCGAGT GAGAGGGGCT
83351 TGAGAGGGAC GACTTAAGAG ATTTAAAGGA TTCCCCCCAG TTGTGGGGAA
83401 CAGAGGACCA GGGTCTGACT CTCCGAGACA CCCAACAAGG AGAGAACTGT
83451 GGCTGGCTTA CTGTTCCCAC AGAAAGAGCT ATACTGTAAT GGTGTCTGTG
83501 TTGAATTTGC CTACACTCCA GCCTTTGTCA GGGCAAAGGA AATGTATTTC
83551 CTGCAGATTA GAGGTGGGCC TAGCGGAGGA GAAGGGTAGC CTGGAGTTTG
83601 ACTCCTGTCC AAATAGATGC CTGAAAGGGC AGTGAGGTTC CAACGGTGAG
83651 TCTTTTGAAT GACAGCCAAA AAGCCAGGAG AAGAATGAAT TATCCAAGTC
83701 AATGGTGCTA TGGCCAGAAG GAACTCTGAG GTGAGACTCT GAATAACCCA
83751 TGAAAGTGCA TCTGAGAAAA AAGAATTAGC TTCAAACATC AGCAGACAGG
83801 TGTTTTGGCA GGAGGGCTCA TGTTGGAATC TCCTTGCCTT CCTTCCCCAG
83851 CCTCCATCTC CCATGCCTAC CTTGGAAAGA TCCGACAGCT AGGGTTGGTG
83901 TAAAGTCGTG GAAGAGAAGA AAAGCAGCTG ATTTCAATCC CTTCCCAGGT
83951 TCCCTGGGGC TGGGGAAAAG ATTTGATTAC CTAGTGAATA TTGGTTTGTT
84001 ACCTAAAGTG ACCATAAGCC TTTCTTTAAC ATTGACCAAA AGAATCAAAT
84051 GGGCCTGTTG AAGTGTTCAT CTAGTGTCAA GGGAAAATTT TTCCCCACTG
84101 AATAAATTTT AAGAAGGCAG TCAAGACAAG AAGCTATATT TGATTATATC
84151 CTGTTAGTGC TTATTCAATA GACACATAAA TCTGTAATTT TTAATATTTG
84201 GTATAGAAGT AGGTTGAAAT CCACAGTAAT TCACAGAAAC TTGTGCAAGG
84251 GTTTTGTTTT CTTTTCTTTT CTTTCTTTTT TTTTTTTTTT TTTGAGACAG
84301 AATCTCACTC TGTCCCCCAG GTTGGAGTAC AGTAGGATGA CCTCGGCTCA
84351 CTGCAGCCTC CACCTGCCAA GGTTCAAGCA ATTCTTGTGC CTCAGCCTCC
84401 TGAGTAGCTG GGATTACAGG CATGAGCCAA CACGGGCGGC TAATTTTTGT
84451 ATTTTTAGTA GAGACGAGGT GTCTCCATGT TGGCCAGGCT GGTCTTGATC
84501 CTGACCTCAG GTGATCTGCC TGCCTTGATC TCCCAAAGTG CTGGGATTAC
84551 AGGTGTGAGC CACCATGCCC GGCCAAGGTC TTTTTTCTTG AAAATATCTT
84601 CACTCATATA AGCAGTATAT GCAATATAAG GATATGCTCT TGGGTTTCTT
84651 GATGTGGTCT AATATTTAGT GTTGGCCCCT TAATTATAAA AGTTGCTTTT
84701 ATCTAAAAGT AGATGTTAGT TGTCAGGCAA TGTTTGCTGT AAAAAATAAA
84751 TAAAAAATAA AAGTAGATGT TAGGATATTT TCTTCTAGCC TGCCTAGTAT
84801 TTATATTAGA TTCTTTCTTT TTTTGAGAAA GCGTCTCGCT CAGCTGCCCA
84851 GGCTGGAGTG CGCAGTGGCG CGATCTCGGC TCACTGCAAC CACCATCTCC
84901 CGGGTTCAAG TGATTCTCCC ATCTCAGCCT CCTGAGTAGC TGGGATTACA
84951 GGCACCCACC ACCACGCCTG GCTAATTTTT GTATTTTAGT AGAGGCGGAG

FIGURE 3-25

85001 TTTCACCATG TTGGCCAGGC TGGTCTGGAA CTCCTGACCT CAGGTGATCA
85051 GCCCACCTCG GCCTCCCAAA GTGCTAGGAT TACAGATGTG AGCCACCGCA
85101 CCCAGCCTAG ATTGTTTCTT AAACCATAGA TGTCTGAACT TTTTTGAATG
85151 AAATTAAATG ATTAGAGATT AGTAAAATTT TTGTATAAGA TAGTAGACTA
85201 ACAAATTCTT ACTAGTCTGG GTGTGGTGAC TCATGCCTGT AATGCCAGCA
85251 ATTTGGGAGG CCAGGGTGGG CCGACCACTT AAGCCCAGGA GTTTGAGACC
85301 AGCCTGGGCA ACATGGCGAA ACCTTGTCTC TACAAAAGAT ACAAAAATTA
85351 GCTGGTGTGG TGGCACACAC CTGTAGTCCC AGCTACTCAG GAGGCTTAGG
85401 TGGGAGGATG GCTTGAGCCT AGAAGGCAGA GGTTGCAGTG AGCAAACATT
85451 GCGTCACCGC ACTCCAGCCT GGGTGACACA GCGAGACCCT GTCTCATTAA
85501 AAAAAGAAAA TTTAACCTGT CTCAAGCTCT CCATACTGTA AGGCTCTGCA
85551 TGTCTTGATT GGATTGTGCT AATATATTTG GCCAATCAGC TCCTTCCTAC
85601 TGTCTACTTT TGAATCCCTG TCACCACCAT CTAAGTCAAG ATGACAGTGT
85651 TTACACAGTC TCTTCATTGT GTTTTAAGAT TATAGTCTTC TTTCTGGTGA
85701 GGGAAGAAAG AAAATAGAAA TATGGCTTAC TGATTGGGCC ATGGCTTACG
85751 CCTGCAATCC CAGCATTTTA GAGGCCAAGG TGGGAGGATT GCTGGAAGCC
85801 AGGAGTTCAA GACCAGCTTG AGTAGCAAAG TGAGACCCTG TCTGTACAAA
85851 AGAAACACAC ACAAAAGAAA TATGACTGAC TAAAATACAT ATAATTTTCA
85901 TAATACTTTA AAATGTAAGA AGGCAAAAAA TTTCTGGGCT CAAGGTGGGT
85951 GATCGCTTGA ACCTAGGAGT TCAAGACCAG CCTGGGCAAC CTGGCAAAAC
86001 CTTGTTTCTA CAAAAAGTAC AAAAATTAGC CAGGCATGGT GGTGCACACC
86051 TGTGGTTCTA GCTACTTGGA AGATTGAGGT GGGAAATTTG CTTGAGCCTG
86101 GGCTGTCGAG ATCACAGTGA GCTGAGATTG CACCACTGCA CTCCAGCCTG
86151 GGCAGCGGAG TGAGACCTTT TCTCAAAAAA AAAAAAAAAA AAAGGCAAAA
86201 AATTAAATTA TTAGTATGGT AAAGTTTCGT TTGGACTTAA TATGAAACTC
86251 ATTTCTAGAA ATGATGATCA TTTGCATAGG GCTTAACTTC CTTTGCTAAG
86301 AAAATAGAGT AGTATACTAG GAGACTTCCA GAGCTGCATA GAGCTTCAGG
86351 GTCATCTACC AAGACAGACA ATTTGTTGTC ATCATCAGTG TTAAACTCTA
86401 AATTATTAAG TGCTTATGTG CCAGATACTG AAGTTTATAT ACACTTTCTC
86451 TAATCTTTAA TAATTCTAGA AAGGTATGTG TTTGATCCAT TTTCAAGATA
86501 AGAAAACTCA TTCAGAGGGG AAGTAACCTG ACCAAGAATA CGTTGTAGTT
86551 GTAGAGCTGG GAATATGACT CATGTCTGTT TTTTCCATAG CCCATTTTCA
86601 CTTAGGTTCC ATTAGTTTAT TATATATTTT AGAAGGACTT ACAGACTTGA
86651 TTTCCTCAGC GAAGTCAGAT TACTCATTTC TTTGGCACAG AATTATGATT
86701 ATTTGTATTT AATATTTCTT TCCTGTTGGT TTTCATTTTT GAGTTTTTAT
86751 GGAATCTGAC TTCTGATTCT CAATTTTTTT ATTGTGAAAT ATAACATACC
86801 TTCAGAAAAA AAGATTAAAC ATATCTATGG TTTTATAAAT GATTATAAAA
86851 TAAATACCCA GTAACTATTA TCCAGGTCAG CAAATATTCT ACTAGTGTAT
86901 GAGTCAATTT CCATGGCAAA AGAACTAAGC TTAGGCACTA TACTCAAAAA
86951 AACTAAAAAT AAAATTTTTC TAAATGTGTA TTATATCAAT GGAATAAATA
87001 CAAATATAAC TTACCATGTC ATAATTCCCC CCACGCTTTC CCTTCTTTTA
87051 CAGCATGGGT AGGTTCTCTC TCCATGGGGA TGATTTTCTT TTGCTGCCCA
87101 ATAGTCAGCG TCTTCACAGA CCTATTTGGT TGTCGGAAAA CAGCTGTCGT
87151 GGGTGCTGCT GTTGGATTTG TTGGGCTCAT GTCCAGTTCT TTTGTAAGGT
87201 AAGGACTTGG TTTTTTCATG TTGCTTTTTA AAAACTGTTA GATACCTTAA
87251 AGTTTTACTT TCAGAAACTA TGCTATTTAC AAGCAAAGAT CCTCCTTTTC
87301 ATTTTTTAAA ACTTTAAGCA ATATGACTTA TAAAACAAAC TGTTATCCAT
87351 AGCAGCAAAT CAGAGCTTGA GAATTTGAAT GCTTTTTTTT CTTGTAATGC
87401 CTAAGACTTA GACATCAATC AGGATATATG TGTTTCTTGG TGCATGGAAA
87451 AATGTTTCTC CTTATCTTTT TCTTCTATTC ACATAAAAAT CCTAATGGCA
87501 CTCAAGTTAT AACAATGATT TATCTAAAAC AGCAGCCTTA GTTTAGGGTC
87551 AGTTCAGTCT GAGGCCCACG GGTTACAATA AGTGGTGTTT TAAAGTAGCT
87601 GTTCATAGGC TTGATATGAA ATATTTTTGT TAATGAGACC AAAACTTTGC
87651 CATTTATTCC AACCCAGGTA GAGAATTCCT GTCTGTTCTT TAAAAAAAGA
87701 ACATGCTAAA ATTTTAAAAT ATCATGGCAA AATGAAGTGG TCCAATGTAC
87751 CTTAAAATAA AACTTAATGT CAATGTACTT CTCCTGTATC TATTAGAATA
87801 AGGATCCCCA ACCCCTGTGC CACAGACTGG TACGGGTCCA TGGCCTGTTA
87851 GGAATTGGGC TGCACAGCAG GAGGTGAGCT GTGGGTGAGT AAGCAAAGCT
87901 TCCTCTATAT TTATAGCTGC TCCCCATTAC TTGCATTACC ACCTGAGCTC
87951 CACCTCCTGT CAGATCAGTG GCAGCATTAG ATTCTTCTAG GAGTGCAAAC
88001 CCTATTGTGA ACTGTACATA TGAGGGATCT AGGTTGCACT CCCCTTATGA
88051 GAATCTAAAT GCCTGATGAT CTGTCACTTT CTCCCGTCAC ACCCAGATGT
88101 GACTGTCTAG TTTCAGGAAA ACAAGCCCAG GATTCCCACT GATTCTACAT
88151 TATGATGAGT TGTATAATTA TTTCATTATA TATTACAATG TAATAATAAT
88201 AGAAATAAAG TGCACAATAA ATGTAATGCA CTTGAATCAT CCCTAACCGC
88251 CCTCCCCCGG CACCATCCAT GGAAAAATTC TTCCTGAAAC TGGTCTCTGG
88301 TGCCAAAAAG ATTGGGGATT CCTGTATTAG AAGAAAAGAC ACAGTGCACT
88351 GAAGAGGGAC TCACTGTAAG GAACGGATGG GCACATAACT GCATGGAACA

FIGURE 3-26

88401 TCTCTCCCTG CAGCTCTAGT TCTCCTTGTA AGTCCCTTGC GTTGGTAGAA
88451 TAATCCTCAG TTAGACAAAC ACTGATTTAA TATGTAGCTC TGGCTAACAG
88501 GAGGTGATTA AGAAGAAAAC CTCTTAAGAT GATTTCCATC CTTTGTCTCT
88551 ACTTTAGTGG TTTATCTTCA TTTCCTGCCT CTTTCCTGTC CCTAGCTGTC
88601 TTTATGCTGC TTTAGTTGAA AAGCGTTAAT GTGGTCATTA AGGAAAAATA
88651 AGTCAAATTT ACATTTGACT TTTTATTTTT AAATATTTAA TCAACAGAAT
88701 CCTTGGTTTT ACTCATTGCC GCCCCCCACC CCCCAACACA CATCCCTTCC
88751 TCAACTCTAA AGTGAGCCTC ATTCTTTCAT ATTTTCTTCC ATCTAATTTA
88801 GAAATTCTAT TTGGATTTTT AAAAATTATA TTTATTTCTT TGTAGAAAAT
88851 AGATATTTTC ATCTTTAAAG TACCTTTATG GGTTTTTTCT TCTAAAATTG
88901 TTTTTTAAAG AAAAAAGTTT TATTTGGAAT AAGATTCTTG TAGGTAATTC
88951 CATGAGATGA TTTATTTTAG CAGCAACATA ATATTTACAT TATTATTAAT
89001 GTAATTAATG TTATTAATAC CTCATCAGAT AGCTTCTTTG ATCTGGAAGC
89051 TTCCAGGTAC CTATTGTCAG TACTTGTGGC TCTACCACTT GCCGAATGTA
89101 TTACAACTCT AGTTGTGGTA GAGAGGGGAC TGAGAGGTAG ACAACTTATG
89151 TAATCTACTA CCTAGTTTGT TAACAAAACA CACATACAAA GCAATGTTTT
89201 TCAAATTTTT CTGACCACTG AGCAATAAAA ATTATGACAT ATATTTTGAT
89251 GTGACCCAGT TCTGTCTCTC TTTCTCTACC CTCTAAGTGA AACAAAATTT
89301 ATTGAAACCA AAATTCCCTT ACTACATGTA ATATTCTCAT ATATTCTATT
89351 AAATTTCGTT ATTTAGCTTG CTGATCAAAG GCTACTGAAA CTTGAGAGCA
89401 AGATACAGGA GCAAGGGGAA ATGTGGTATA GATTCTGAGT GTCAAGTGGC
89451 AGGTCCATTT TTTCCTCTAG CTCCAGTTCT GCCTTCTGAG GAAAACCTTC
89501 TCCAACAACT TAGGTCAATC ACACCCATGT CCCTTCTCTG AATCCTTTTT
89551 GCACATATGA TTGGTATCCG ACAGCCTTAC TCATTTACAT TGCACTTATT
89601 TGGCTGCCAA ACGTCACAAA CTGGAACCAT GTGTTACTGA AGGGAAAACC
89651 TGGAAGTGAA AAGGGTTCAG CAGTAGTGCA AATACCATCA TAAAGCTCAT
89701 ATACTTCACT CTGCAGGAGG GAGAAGCTCT GTGGTTTTCC AACTGAGAGC
89751 ATTACAGTAC AGTGATACCA CTGTACAGGA ACTGATGTTC CTGATGATTC
89801 TGCTGTGAAC AGTATTTTTA ATATACACTT TGAAGAAGGC AGAGAGAAAT
89851 GTATAATAGA CTTAAATTTT TTTCTTTAAA ATTGTTAAAT AAAAACAAAT
89901 AAGCACTTTA AGTAAGTTAC AATTATCTGG AAAACTACTT AGGTGGAAAA
89951 ACTGATACAG AATGAATGAA GTATTAATTT CTGTTTTGTT CTGTGTTATT
90001 ATTATTTGGG ATAGATGTCT TGTTTCTTTA AGCAGACTAT GAATATCTTG
90051 AAGGCAGAAC CACATTTTTT TTTTTTTTGA GACAGGGTCT CACTATTACT
90101 CAGGCCAGAA TGCAGTGGTG TTATCATAGC TGACTGCAGC CTGGATTCCT
90151 GGGTTCAAGC CGTCCTCCTG CCCCAGCTTC CTGAGTAGCT AGGACTACAG
90201 GCATGTGCCA TCACACCCAG CTAATTTCAG CTATTTTTTT TTTTTTTAAA
90251 TAGAGATGGG GTTTTGCTAT GTTGCCAGAC TGGTCTCAAG CCATCCTCCT
90301 GCCTTGGCCA CCCAAAGTGT TGGGATTACA GGTGTGAGCC ACCACGTCTG
90351 GCCAAGGACC AGATTTTTAA TATTCTTTTC CACAATGTAT CTGGTACACA
90401 GTAGTTGCTT AATATGTTGG CTAAACAAAG AGTGGAGATT CAGTAAAGGG
90451 TGATCAGAGT GAGGTGAGAT TAATTTGGGA AAGCCTAGAA GTGATTCTTG
90501 AGCCTGATTT GAAGGTGGTG CTAGCTGTGG ATTAATAGAG GGAGAAGGGC
90551 ATCTCAGAGA GAGGATTGCC AACATGCCTT AATTTTATCA GATTCTAGAG
90601 TTCCTTATGA TTACCTCAGC ATGTTGCTAG ACTAGCATTA TTATCCAAAA
90651 TTTTAATTAT TAACCAACTT TAATCTTACT TTCTAACAAA TTGTTTGCTT
90701 TTACTACTGA TAGCCTTTTC AAAAAACTTT AACTAGTTTT ATTCCTTACC
90751 ATAATTGTTT CAAAGAACAT AATGATATGA TCCTTTATCT TCCTAAGAAA
90801 TGTGCAATTA TTTGGTTAAA CTGTAAGATT ATTTAATCCA TTATTCTTTT
90851 GACACATGCA TGGCCTTACA GCTTACAAAC TGGGATCACT AAAGGAATAC
90901 ACTTAATTTA AGTCTTTCTG TAGTCAGAAT ATGATTTCTT GTTGTCTTGC
90951 ACAATACTGA GAACAGTGCA GTACAGGGCG AAGGTTGGTC TACAGCCCTT
91001 AGGCCAGCAA AAACAGGCAC AACTGCACCT CTGTGCAAAT GTTCCTGACA
91051 TAACCTTGGG GAAAAAATAT AAAATGCGGC CTTTTCTTTT ACTACCTTGT
91101 TTGGTAAGTA CCTGGAAAAA CTCCATGAAA TAATTAGATT TCATAGTTAA
91151 TTCTAACTTT TTTAAAAAAT GTTTCATTGA GACTAGGTTT TTGGTTTGTT
91201 AATTGAATCA CTGTTGATTT TACCCTTCCT GGCACCAACC TTTATTTCTG
91251 AGCTGTGGAG AGCACAGTTC TCACTCAGTG CTGTGTGCGT CACCTGAAAT
91301 CCACAGAAAG AGGTGGCTGA ACAAAATCAC TGATGACCTT AATGGTTATT
91351 TTTCACATAT TCAGATTAAA TTAAAATACG TTTAGTGCTA CATGCTTGAC
91401 TTACTGAGTT TTTCCCTCTA TTTTGGTTAA TTTTTTTTTT TTTTGGTTAA
91451 CTTTTACTTG TAGAAAATAT GTTGATGAAC AAAAACCCAC TTATACTATA
91501 AGATTTTATT CTACCAAGCA CACAGTAACA ATATTGAAAG CTGCTTTCCA
91551 TCTTTTTCAT CTTTATACAG TTCCATCGAG CCTCTGTACC TTACCTATGG
91601 AATCATATTT GCCTGCGGCT GCTCCTTTGC ATACCAGCCT TCATTGGTCA
91651 TTTTGGGACA CTATTTCAAG AAGCGCCTTG GACTGGTGAA TGGCATTGTC
91701 ACTGCTGGCA GCAGTGTCTT CACAATCCTG CTGCCTTTGC TCTTAAGGGT
91751 TCTGATTGAC AGCGTGGGCC TCTTTTACAC ATTGAGGGTG CTCTGCATCT

FIGURE 3-27

```
91801 TCATGTTTGT TCTCTTTCTG GCTGGCTTTA CTTACCGACC TCTTGCTACC
91851 AGTACCAAAG ATAAAGAGAG TGGAGGTAGC GGATCCTCCC TCTTTTCCAG
91901 GAAAAAGTTC AGTCCTCCAA AAAAAATTTT CAATTTTGCC ATCTTCAAGG
91951 TGACAGCTTA TGCAGTGTGG GCAGTTGGAA TACCACTTGC ACTTTTTGGA
92001 TACTTTGTGC CTTATGTTCA CTTGGTGAGT ATGCTCCTTC ACTGATCATG
92051 AATATTACTA TTTAATAAAG AAAAACTTCT TTGAAGAGAA AGTTAGGTCG
92101 AGTTAAAGTT GGCCTCAAAC ATTATCCTGG TTGTAATTTT GGTATTCTTG
92151 AAATGAAAGG TCTCTCAAGA CAATGTCAGC ACATCCATTA GACCACTAAA
92201 CAGAGAGAGT ATGTTTCATA GTGTGCTTTG GTATTTTAAA AACCCTGCAA
92251 ACCCAGCCAG ACACCATGGT GCCTGTCTAT GGTCCCAGCT ACTAAGCTGA
92301 GGCAGGAGGA TCACTTGAGC CCAGGAGTTC GAATCCAGCC TAGACAACAT
92351 AGAGAGACTC TACCTCTAAA AATAAAATAA ATGTCCCCAA ACAAACACAA
92401 TGTTTTTTAA CAGGAAGGCT AAAATAGTGG AACAAATTAC AATCAGTATA
92451 AAACATTTGA TAGGTCTCTT TTTCTTCATA TGGCTTTTAT CAGGGACAAA
92501 GCTAGCGCTA TGATTTTGCT ACCATAAGTA AATTGTTTTT CAACCGAAGG
92551 GTGTAGGTAA TTAGCAAAAA AGCCATGATG TTGATACAAA GAAACATTAC
92601 ATCTACTTGT GGTACACTTC TGGGAAAATG GGAATTCTAT TCAGAGGAAT
92651 ATCTGAGAAA AGTTACTCAA GATCTAAATG AGGAAAGAGA ACTATGGTTT
92701 TATAGGAAAT TAGGATTTCA AGTGCTCAAG AAGTTTATAT TGTTTATTTC
92751 TATTTCAAAG GCAAAATTCA GCTTTGTTAT ACTGAAATAC GAATAATTAA
92801 TGTCTAGACT GGGGTTGGTG CCTCACGCCT GTAATCCCAC CACTTTGGGG
92851 GGCTGATGCA GGAGTTCAAG ACCAGCGTGG GCAACATAAG GAGACTTCAT
92901 CCCTACCTGG GGAAGGAAAA AAAAAAAAGA AGGAAGAAGC AGTGTCTAAA
92951 GTATCTGCCC CTGGCAACGT TTGTTCAAAA GTGTTCATTA TGTTTCTTCC
93001 TTTTTTCTTT TGTGGCTGAA AATGTATTTA CAATTCACCG TAAATGATAA
93051 AAATGGCATT GGCACACATA TTTGTATGTT TGTGAACTTG GATTTTTTTC
93101 TAGCTTACAG TCTACTTTTG GAGATTTGTG CAATTTTTCT TTAGTTAAGA
93151 AATAAGTATA AATATAACCG ATTTACGGAC TATCAGGCTA CATCCTGATC
93201 TGATAGTCCA TTTTCATACT ATTAGGAAAG TATAGCCGAA CCAACTTAAG
93251 GTAAGTTTCC TGGAATATAG ATCTGTTGTG ACAGGATTAA CTTTACCATC
93301 CAACCTCTTT CATAGCTTCT GTAGTCAAGA GAACATTTAT TGTGTCCTTT
93351 CTTAAAAAGA TGAGTAGAAA TTCTTTTTCT TTTTTTTCTT TTTTCCAGAC
93401 AGGGTCTTGT TAAGTTGCTC AGGCTGGCTT CAAGCAACCC TCCTGCCTCA
93451 GCTAGGATTA CAGGTGCAAG CCACCACACC CAGCTTTAAA AAAAAAATTC
93501 TCTTTGGTAC TACCACATGA ACACACCTAG AGAAATCATA ACTCAGCTTT
93551 GCTAATACTA GACATTTACC AAAGGAAAAG TGGTAGATGA CTGTCTAGTT
93601 ATTTTTGGTT ATATATTTAT AATTTGTAAA TTAATTTCAC ATATATTACT
93651 TCATTTGACT TTCACAATAA ACCAGTAAAG CAGATAAAAT AAATATTAGC
93701 TCCAATTTTA CAGACTGAAA AACAGATCTA TTGTTAATAG AGACGTTAAG
93751 TGATTTTCCA AGAATTACAT GTCAGTAAAC AGCAGAGCAG GAGTTAGTTC
93801 TCTTCCACTG TGCTTACCTG GTAGCAAAAT CAGTCTACAG TCTTAATAGC
93851 ATATTGGGCC ACTTCCCTGG ATATATTACC AAATGTGTCC ATCTTATTAG
93901 GGGAAAAATG AGTATGCCTA AGGAAATTTA ATAAGCATGT TATTTCTTCA
93951 GGTAAATAAA ATTTCATAGT GGAAGGTGAG TTAGACAATG TTATAGATAC
94001 TTTTGTGATC AGGAGATGGC AAATCAGATG GTGCACAGAA CAATAAAGTC
94051 TCTGTTAATT CTGTTAATAA ACCATGCCTT TTTTCTGCTT TCCCTTCTTC
94101 CAGGCATGTT TTCTTACAAA ATATGTTGAC ATTGTTCATT TGAGATTTTC
94151 TCTTTCTCAT AACGGTGCCC GTTATCGCAC CGAATGCAGC ACGGTAGAGG
94201 AAAGATCAGA TAGCTAAATG CCATACAGGT GTTTAAATCT CCTCTTTGGT
94251 TATGTACTGA GTTTGTCACT TTGTTGTAAT TTAAGGTTTG AATTATGGAT
94301 ACTTAACCAG GAATGGGACA CTAGTTTCCT CCTTATACAG GGAAAAGGTG
94351 TCTCATATCC TTCAAAAGAC TAGTAAAGTA GATGATGTTC AATTCCTACT
94401 AAACCCTTTA TTGACTGTTG AGGGGACACA TATATGAGAC GTAAAAATTT
94451 GCTCTGAAGG AGCATAAACC TAGTACATGT AATTAAAAAT GGCTACAGTT
94501 TATAAAGCAC TTTTACATAC ATTCTCTTAT TTAATATTCA CAACAATGCA
94551 GTACCTGTGG TGTATCCTCT TTATTTCATG GAAGGGAAGA CTAAGGCCCG
94601 GAAAGATTAA ATAACTTGCT CAGCCAGGCA CAGTGGCTCA CGCCTATAAT
94651 CTCACCACTT TGGGAGAATG AAGTGGAAGG ATCACTTGAG CCCAGCAGTT
94701 CAAGACCAGC TTGAGCAACA TAGTGAGATT CCATCTCTAC AAAAAGTAAT
94751 TTAAAAAAAT TATCTGGGCA TGGTGGTGCA TGCCTGTGGT CCCAGCTACT
94801 TGGGAGGCTG GGGTGGAAGG ATCGCATGAG CCCAGGAGGT CAAGGCTGCA
94851 GTGAGCCATG ATGGTGCGAC TGCACTCCAG CCTGGGTGAC TGAGTAAGAC
94901 CCTATCTCTA AAAAAAATTA TAAAGTATTC TAAAGGAAGA ACAGATTGAA
94951 CAATTTTTAA TTTATTTGTC TCCTCCTCCT AGTGGCAGCC TTTTAAATAT
95001 GGAAGGTGAA GAAATAAAGA GCCAGATGTG GTGGTACACA TCTGTAGTCC
95051 TAACTACTCA GGAGGCTGAG GCAGGAGGAT TGCTGGAGCC CAGGAGTTCA
95101 AGGCTGTGGT GTGCTATGAT TGTGCCACTG CACGCCAGCC TGGGTAACAG
95151 AGCAAGACTC TGTCTCTAAA AAACAGATAA TAAATAAAGA AGTAACTTGC
```

FIGURE 3-28

95201 TTGAGGTCAC AGAGATAGTG ACTGATAATT ATTACTGTAG TACTTTTATG
95251 TAAGAGGCAG TATTGTATAG TGGTTTAAAA GTGAAGGTTC TGGGCCTGGT
95301 GCGGTGGCTC ACCCCTGTAA TCCCAGCACT TTGGGAGGCC AAGGCAGGTG
95351 TATCACCAGA GGTCAGGAAT TTGTGACCAG CCTGGCCAAC ATGGTGAAAC
95401 CCTGTCTCTA CTAAAAGTAC AAAAATTAGC TGGACGTGAT TGCTTGCACC
95451 TGTAATCTCA GCTACTCAGG AGGCTGAGGC AGGAGAATCG CTTGAACTTG
95501 GGAGGCAGAG GTTGCAGTGA GCTGAGACCG CGCCATTGCA CTCCAGCCTG
95551 GGTGACAAGA GCGAAACTCC ATCTCAAAAA AAAAAAAAAA AAAAGTTCTG
95601 AAGTAAGACA GATCTGGATT TAAATTCAGG TTTTGTTTCT TACTAGTTGC
95651 ATAACCTTGG GCATCCTCTG TAAGCATCAG TTTCCTCATC TATGGAGATA
95701 AACCCAATTT TGCAGAGTTG TGAGGATTAG ATAAAATGTA TGTGAAACAT
95751 CTACCTCAGT TCTGGCATAA AAATGGGAGT TATTTTAATG TAAGGCAATG
95801 TGATTGCCAA CTTGAGATAG AAGTAAATTT TGAAAGGAGA AAGATAATAC
95851 CCATTTGGAA AAGTGGTTTT AAAAAGTTTC ATAGCATTGG AGTTGGGCCT
95901 TGAGCATGAG ATTTTGTGTA CAAATCTGAT CTTTGATCAA CTAGGGAACT
95951 AACTTACCAG TTTAGGTCTT TGAAGATTCA GAAATACAAT GGAGTGCTCT
96001 CATTGCTATG TTAAAAATTC TAAGATCTTA TTAGATTGTA CATGATGATT
96051 TGAGAGAGAA TATGTATGCT TGCTTTCAAA GTGAGGTTGG AGGTTTGATC
96101 TTCTCGTAGT TGACGTTTCA AAAAGAAGAA TTAGATTGCC TCCTCGAAGC
96151 TAAATTTACC TTTCTTTTAG GCCTTCCCAC TTAAAATCTT TTTTAGAAGG
96201 ATACAAATCT TATAGATCAA TTTAGATGAG GCCTAACTTT CTAAAAACGA
96251 TTCCTAGTAG CAGCTGCATC AGTTTTTATG AATTGCCCCT TTTGCCTGAG
96301 AGTTGTTTTG TTTTGTTTTC TGGAATCTTT TTTTGTTTTG TTTTGTTTTG
96351 CTTTGTTTTT GTTTTCGTTT TTTTTTGAGA CGGAGTCTTG CTCTGTCTCC
96401 CAGGCTGGAG TGCAGTGGTG CAATCCCGGC TCACTGCAAC CTCTACTTCC
96451 CGGATTCAAG TGATTCTCCT GCCTCAACCT CCCTAGTAGC TGGGATTACA
96501 GGCGCCTGCC ACCACACCTG ACTTAATTTT TTGTATTTTT AGTAGAGACA
96551 GGGTTTTGCC ACATTGGCCA GGCTGGTCCC GAACTCCTGA CCTCAGGTGA
96601 TCCACCCATC TTGGCCTCCC AAAATGCTGG GATTACGGGT GTGAGCCACC
96651 ACGCCTGGCC TCTGGGTTTC TTTTTTTTTT TTTTTTTTTT TTTTCTTTTT
96701 AACGGCTCCT CTGACTCCTC TCATTTAGCT TTCAGGAGCA TAAACTCTCT
96751 TGGTTTTCTG CCTACCTCCA CATCACTCCT CCTTAGTTTC TTTGCTCACT
96801 TCTTCTTTTT GCCACTGACC CCTGAATATC AGCATGTCCT AGGGCTTGTC
96851 CCCTGATCTT TTTCTCCATG TATTCTACTG GTGGTTTCAT CCAGTCTCCT
96901 AAGTTCATAC ATCACGTATA TGTCAATGAC TTCAAATTTA TAATTCTGGT
96951 CCAGACCTTT TCCCTGAATC CTCCACCAGA GCTGTATATC CAGCTGCTTA
97001 CTTAACATCT CCACTTGGGT AACTGCTAGG TGTTTCAGAC TTACCCTGTC
97051 TAACCCTGAG GTCTTGATCT TACCCCTTAA AACTTACTCT GCCCCCAGCC
97101 ATCCTCATCT CAGGAGCTGG CAATTCCGCC CTTTCAGTTG ATCAGACTCA
97151 AAACTTTGGA GTCATCCTTG GCTCTTCTTT CTTGCACACC ATAGTCCTGA
97201 TCCAGTGAAG AAATCCTGGT GGCTTTTCTT TCAAAATATA TCCAGGATCT
97251 GACCACCTCT CACCATCCTC ACTACTCATA CCTTAGCCCA GGCTACCACG
97301 TACCCCTAGC CTGGATCACT GCCAGAGCCT CCTAACTGGT CTCTCTGTTC
97351 CTTCTCTGCC CCGCGGAGTT TGTTCTCTAT GAAGAAGCCA CAGGCATTCT
97401 TTCTAAACAT AAGTCACTCT GCTCAGAATC CTTCAATGGC TTCCCATTTC
97451 CCTAAGAGTA AAAACCAATA TCCTTACAGT GACCTACAAG GTCCTTCACA
97501 ATCTGGCCCC CACTACCTCT CCGAGCTTCC ATCGCTGTCC CTTGCCCACT
97551 CTGCTTCTGC CATTCGCCTT TTAATGGGGC TCACTCTGAC TACCTGCTTG
97601 AAACTTCCTG CGTCCCTTTT CCCCTGAGTA TTCACAAACC GCTCCTAGTA
97651 CTCCTTTTCT TTTTTTGTAG CACTTAATAC TTTCTAACAT TATCTATTTT
97701 ACTTCTTTAT TGTAGTCATT GCTTACTATC CGTATATTTA CACGTCTGCT
97751 AGAATGTAAA CACCACAAGG GTAAGGATCT ATTTCATTCA GTGGTAGATC
97801 CCAAGCATCT AGCACAGTGC CTAGCACACA CTGGGTGCTC AAATATTTGT
97851 TGAATGACTA AATATATTCT GGGTGAGTCT GAAGTGACAC TGTATAAGTA
97901 ATGTTCATTT TTTCATCATT TGGATCTTTA AAATTCTCTA CTTTGATGCT
97951 ATAATGATTT TTCACATTCT GTACTTGCAG GACATGGTGT TATTAATATT
98001 TATTCAATAC TTATTCAACA AATAAGCTCA AACTAAGGAA ACCTCGGAAT
98051 AATTGAGTAA CCAGTAATGC TGTCCGTTGA TGGAGGAGAG AGTTGGTGTG
98101 TTTTGCTCTG ATTCACTTAT GCCTTTGCTG AAATTTTAAG ATAAATAGAA
98151 GAAATTTCTG GTCCCTCAAG TAACTGTGTC TTCAGTACCC ACTGAAAAAT
98201 CTCAAAGAGT CTGGAGTGGT GTGTTTAAGA ATAGGATGCA GGATGCAGAA
98251 CCATAACCAG GCCTCAGGTC TGCATAGCTT TGGTCGAGCA TTGAGCATAG
98301 GGCCTCGTGA GATAACTGAT AAATGCCAAA TATGACAATG ATAAATGCCA
98351 AATATGACAA TGATAAATGC CGAAGAATGA CAGTGACAAT GATAATGAAG
98401 TTACCAAAAA TGATGGTAAC TTTTCTCATT GGCATGAAAT GCTCTATCTC
98451 CAATCTGAAG CTGATGATGT AGTTTCAGTT ACTCTCATCT CTCTCCCCTG
98501 CTACTCAGAT TGAAAATCAG CTACTTAGTA CCTGTGTTCT TTGACTCTAG
98551 ACCATATCAT TGGGTCAAAT TTCAGTTTTT AAATTTTAGA TCCACATGGT

FIGURE 3-29

98601 TCTCTGTCAA GAAGATGACT GACTCATATT GAAATCTGTA AAATATGTAT
98651 TCATTAGCCT GTTTTTTAAA AACTCCCTTA TAAGTGGGTT GACTTTGTGG
98701 CAGATAGTAA TTGACTGTTC TCAAAAGAAA CTTTGACCTG GTAGGAAGAT
98751 CCCATTTACC TGATGCTATG GTTCAAGACA GACAGATCAT TTGCTTGCTA
98801 GCAGGGCAAT TAGGTGAACT TCAAGTCCAC TAGTAATTGG AAATGATTTT
98851 TTTTTTTTTT TGAGACTGAG TCTCATTCTG TCGCCCAGGC TGGAGTGCAG
98901 CGGCATGCTC TCGGCTCACT GCAACCTTCA CCTCCTGGGT TCAAGCGATT
98951 CTTCTGCCTC AGCCTCCCGA GTAGCTGGGA TTACAGGTAC CTGCCACCAC
99001 GCCTGACTAA TTTTTGTATT TTTGGTAGAG ATGGGTTTCA CCATGTTGGC
99051 CAGGTTGGTC TCAAACTCCT GACCTCAGGT GATCTGTCTG CCTCGGCCTC
99101 CCAAAGTGCT GGGTTATAGG CATTAACCAC CGCCCCTGGC CATGAATTGT
99151 ATTTTTAAAC CAGAAATGAA AATTTGAGAC TAATAAGTCA GTACAGGGAG
99201 CATGTAAACC TCGAAAGGTA TTTTTTAGCT TTGAGTAGTG CCAGATGCTG
99251 CCAAGGGTCA ATCAACACTG GAATGTAGCT ATTAGACCTT GCTAGGCAGA
99301 GCACCTCCAT TTACACTGTG GTCAGAGCAG CAGTACTGCT CCAAGCCAGA
99351 GCTAAGGGCG CTGAGCCACG CAAATAGGAA CAGCATACAA GCCTTCATCT
99401 CTCTGTGGCT TCCTCAGAGG GAGATTCATG TAACATTTGC CAAGAATTGA
99451 TTATGTGTCA AGCACTTCCC CAAAATCTCA CAGAACCACC ACAAGGATGA
99501 GTGTAATAAA TAACACATAC TTAGAGCCAA GGAAACAATT CTGCAAAGCT
99551 GTGCTTGTTC AAAGCCATTC GCATTATGCT TAAAGCTGGG ATTTGAACAC
99601 AGGTTTCAGA CAAATATGTC TGAAATATAC TCTTTTTATG AAGGAGTCTG
99651 CATTCCTTCA TTGCTAATCC AGAGATAGGA GTGCTGCTAT TTTCAGCCAT
99701 ACTGGGCCTA CACCAAAGAT TGCTTTGCAC GTTTCCCTTC TGTTCTCTCA
99751 GAACGAAGAA CAGAGGCCAT GTTGAGCTGT TCCAGCGCTC AGAGCATGCT
99801 TCACAGCCAG GGAGAAAACT CTGGAGGAAA CCAGCTTTTG TTTTGATATA
99851 ATTAATGGGA ATGAGAAAAT ATCTATACCC TTATTTTCAG CCCCAACTTC
99901 TCTTTTGATC TCAAGTACAT TGTGAATATG AGAAAACTGA GGCCATGCAG
99951 TTACTTTTCA CAACCTGTGA CAAGCAGAAC ATGGACCATA CATAGCTTTG
100001 TGTTCAATTT TGCTTTCTAC AGTAAACATT AAGCATAACA GAAGAACAAA
100051 AATGGACATG TACAAATTTA TAGCAAGATC TATCCTTTAT TTGATTAACA
100101 TAAATACTAT TGCAGGAAAA TGGAAAAAGG TAAACTGCTT GAAATTTAGT
100151 CACATATAAA CGCTCCGAGG CCACTGGTGG ATCATTAGTC TCCTGAGAGA
100201 GCTCTAAAGA ATTAGTGTGT TGGAAAACTG TTCCCTCCTG TTAATGTGTA
100251 AATTTCACCA GTGGGTTTTT TTTTTTTTTA AGACAGGGTC TCACTCTCTT
100301 GTCCAGGCTG GTATGCAGGG GTGCAATCAC ACCTCACTGT AGCTTCGACC
100351 TCCCGGACTC AAGCAATCCT CCCACCTCAG CCTCCCAAGT AGCTAGGACC
100401 ACAGGTGCAC ACCACCACAC ATGGCTAATT TTTAATTTTT TGTAGAGATG
100451 GGGTTCTCAC CACATTGCCC AGGCTGGTCT CAAACTTCTG TGCTCAAACA
100501 ATCCTCCTGC CTTGGCTCCC TGCCAAAGCA CTGGGATTAC AGGTGCAAGC
100551 CACTACATCT GGCCTTCACT AGGTTTTTCA TTTTGTTTTG CATGTGTCTC
100601 AGGTTTTATT TGATAAAATG CAGTATACTT TGAATCATCT CAAATTTCAT
100651 TTCTAATATG GACATTGGCA TGTCTCAAAT CCTTGGACTA ATAATCAAAT
100701 TAAAGTTTGT TCAAGTTTGA GGAACCTAAA TTAGCCAATT AGATAAGGGT
100751 CCTTTCATGT TTTTATATCA ACTAGAAAAT AAATTGTTTT GATATGGGAT
100801 GAATAGAAAT AGAAATCTTA ATTTGAAGAA TCTTCCCCTT GTGAGGCTAT
100851 ACTAAATGGC TTTTGCCTGA TATTTACAAG GTGGCTTTGG GTTGTGGAGA
100901 GAGTTGTCTG ATCCATTGAG AGTACATTTC TTACCTTCAA CATCTAGGGC
100951 ATCCTTTGGG AGAAGCCCTT GTAGTCACTA ACTCTAAGGA TCATAGAGCA
101001 TAAGGGTAAG CAGGCCTTCT TATGTATTCA TGCTATCAGG AAGGGTCTTT
101051 AGCACCCCAAA CAAAGTTCTA GGGGCTGTAC ATTGCTGATG TGTTAACCCT
101101 CAGCTGCCCA TGTAGCATCT ATTTACCCCT ATGCTTTCCC CACTTTCTAT
101151 CCCTATCATT ATATCTCTGG CTCTTTTGCC CCTCTCTCCT TGGGCAGCTT
101201 ACTTGTAATT AGAAAGTTTA TATTCCCTCA TAACATATTG TAAAAGTGCT
101251 CATTTAAAGG GCAATGCACA CCAAATTGGA GGTGTATAAT TGCAAACATG
101301 GAATCCCTAT ATCTCTGTTA TGCAATCCCT GTATCTCTGT ATCCATGTTA
101351 AATTGAACTG ATGCTTTTTT GAAGTAAAAT GGTAAGAACA GTGGCAACAT
101401 CTAGTCTTCA GAGCATAGTT TAAGATTTTT GCCCAATCCT CCAACCCATG
101451 CAATGGTGTG CTTTGAAAAC CACAGGTTTC TTTTAGACAA ATACAACATT
101501 TATTTCCCGC ATTTCTTTTT GATTTAACAT TTTAGTTAAC ATTTTTATTA
101551 ACATTTTAGT CTACAAGATG CTTCCACATT ATCTCTCATT GGAGTCTCAG
101601 GACCACTGTG TGAAATGGGC AATATCAGGG CTTCTATCTA GCAAGAAAAG
101651 AACCAGATTT GGGGTGGCGA AACAACTTGC TCAGGGTTGC AAGGATGGTA
101701 CATGGTGCAG CCAGGGCTTG AGCTTGGGTC TTCTTAAAAG TGTGGCTTTT
101751 AAATAAAATA CTTAAGTGCC TGCCAAAAAA GTATAACATT AACTTAGGAC
101801 CTGAAAGGCA TTGTACAGAT CAGGTAGTTG CACTCCTCCC CCTGCCCTAC
101851 AAAAAAAGAA AGGTAAAGGA ACGAAGGCAT GGAATAGTTA AGTTGCTTGC
101901 CAAAAGCCAC AGTTATAAAA GTAGCAGAAC TGGGTGTAAT ACCCAAGAAC
101951 ATCCATGGAA AATAAATGGA AGCTTATTAC AGCCCAGCCT GTAAATATGT

FIGURE 3-30

```
102001 ACATAGAAAC AGAATGTGTA TGTAGAAACA AAATTATTAG AGGAGTGAAA
102051 TTAGTTTCTG TCCTAACCCT GGTCAACTCA TAGGGTTCAT TTCCCAATCA
102101 GGAGGTTGCT TTGCCATTTT GACATGATCC ATTTTCTCAG AGTTGCCAAG
102151 AGCAATTGCA ACTGGCTTTG TTGTCTTTCT TTCATGGATC TATTTGGAGC
102201 TCAGGGACAA GGTTGTTATG CGGCAGATTG TCCCTGACAA TGCAAGGCTC
102251 ACAGTAGTAT CCTTATAAAC CAAAGTTCAC TGTACACTGG GCCAGGAATG
102301 GGCTGCTGAA GGTGACCTCT CCTTCATGTT GGTTGGTTTA TCTTTGTTGG
102351 TTGGTTGTGT TTATCTCTGT TCACTATAAG GTTCTGACAG AAGCAAAGTC
102401 TTGGTCCCCG TTACATGTCC CCAGCCTGGT GGATGATGGT TACAGAGTGC
102451 TGGAAACTTT TTTATTTATT TAAAATGGAG TTACGTAACT GAAGAGTTGC
102501 CCTATCTCAG TCAGCACTGC AACCAATTAA AATAGAAAGG CTTAAAATAT
102551 TAATTTTGTT TTTAGCCTAG AGTCTTAAAT ACTAGGGTTT AAAAGTTTCA
102601 TTTACTTTCT CTCTCCCTCT CTCCCTCTCT CCCTCTCTCC CTCCCTTCCT
102651 TCTTTCTTCC CTCCCTCCCT TCTTTCCTTT TTCCTTTTTC TTAACTATGC
102701 AGAATGCTTT GCTACTTTGC AGAATGCTTT GCTACTTTGC AAAATAAAGA
102751 GTAATCACTT TGCTTTTTCT TCAAGACTTT CAAATTTAAG TAATTTTGTT
102801 TTTCTATTTT TCATTCAAAA ATAGGGTTAT AACATTTCCT TGACACAAGG
102851 ATATATATAT ATATTTAGAG ACAGGGTCTT GCTATGTTGC TTAGGCTGGC
102901 CTTGAACTCC TGGTGTCAAG TGATCCTCCT GTCTCAGCCT TCCCAAGTGC
102951 TGGGATCACA GGCATGAGCC ACTGTACTCA GCCTTTATTA ATCTAAATAT
103001 GATAATTTAC CCACTGAGAT TCATTTGTGC TGATTAAATT TACTCTCAAT
103051 CCCTATATGT ATTTCTTGTA TTTTTGTTGT TGTTATGTGG CCTGGAAATG
103101 TTTCCTTTAT TCATTCATTC ATTCCCTTCG CAATTAACTT CCAAAAAGGC
103151 TATGAAGATA TTTATGCACA TATACATTTT ATGATTCAAC TCTCATGGTA
103201 ACCTTTTATA AGGAAAGCAG TGCTAATAGT GGTTCTTCTA CTAATAGATA
103251 AGAACATTGA GGCTCAGAGA GGTTTAAAAG GTTTGCCTAA GGTTGCACAG
103301 TTAAGTAACA GAGTTGCCAT TAGAATAGAT TTCCAGTTTA TTCTAATGGC
103351 AACCTTAATT ATTGTTGTTT GACTTGCTCT AATTATAAAT CAGTTAAACA
103401 GATTCAGGCA TTCATTTGAG TATCTATTGT GTGCCATAGA CATTCTAGAT
103451 GTGGGGATAC AATGCTAAGA ACAGATACTT TCTCATACAC AGCTCAATAA
103501 ACAAGTAAAT TCATAAAGAA ACAAAGTAGT TTTCAAATAG TGATATTTTC
103551 TTCAAAGGAA ACAAATTGAG ATGGAAGGTG ATGGTAAGAG ATGGTAAGAG
103601 ACAACTTTGG CCAGTCAGGG ACAACCTCAT TGAAAAGCTG ATAGTAAGTA
103651 CATCCTTTGG GTAAAGGGTA GTATAAGGTA CTTTGAAGGT ACAAAAATAA
103701 GACAGCTTTC TATTGCCCTT GGGAGGCCTA TAACAGAATT TCTCAAGTCT
103751 CTAAGGCCAA TCAAGAGTTG GATTTTTTTA TCCAACTTAT TTTTAATTGA
103801 TGTATTATTA AAAATCTGCA TATCAAAAAT GAAAATGTCT TGCATACTTT
103851 GCTGTAGGAC CCAATCATTG TTTTTTCTTC ATATACTGCA TTAATCTGTT
103901 TTCACACTGC TAATAAAGAC TTACCTGAGA CCAGGTAATT TAGGAAGAAA
103951 AAGAGGTTTA ATGGACTTAA AGTTCCACAT GGCTGGGTAG GCTTCACAGT
104001 CATGGTGGAA GATGGAGGAG GATCAAAGGC ATGTCTTACA TGGTGGCAGG
104051 CAGGGGAGTA TGTGCAGGGG AACTGCCCTT TATAAAACCA TCAGATCACA
104101 TGAGACTTAT TCACTGTCAC GAGAATAGCA CAAGAAAAAC CTGTCCCCAT
104151 GATTTAATTA CCTCCCACCA GCTTGCTCCC ATGATATGTG GGGATTATGG
104201 GAGCTACAAT TCAAGATGAA ATTTGGGTAG GGAACACAGC CAAACCATAT
104251 CATTCTGCCC CTGGCTCCTC CCAAATTTCA TGGCCTCACA TTTCAAAACT
104301 AATTATGCCT TCCTAATAGT CCCTCAAAGT CTTAACTCAT TTCAGCATTA
104351 ACTCAAAAGT CCACAGTCCA AAGTCTCATC TGAGACAAGG CAAGTCCCCT
104401 CTGCCTATGA ATCTGTGAAA TCGAAAGCAA GTTAGTTACT TCCTAGATAC
104451 AATAGGGGTA CATGCATTGG GTAAATACAC CCATTCCAAA TGGGAGACAT
104501 TGGCCAAAAT AAAGGGGCTA CAGGCCCCAT GCAAGTCTGA AATCCAATAG
104551 GGCAGTCATT AAACCTTAAA GTTCCAAAAC GATCTTCTTT GATTCCATGT
104601 CTCACATTCA GGGCACATTG ATGTAAGAGG TGTTCTCCCA TGGCCTTGGG
104651 AAGCTCTGCC CCTGTGGCTT TGCAGGGTAC AACCCCCCTT CTGGCTGCTT
104701 TCATGGGCTG GCATTGAGTG TCTGCAGGTT TTCCAGATGC ACAGTGTAGG
104751 CTGTCAGTGG ATCTACCATT CTGGGGTCTG GAGGACGGTA GCCCTCTTCT
104801 CATAGCTCCA CTAGGCAGTG CCCCACTGGG GACTCTGTGT AGGGGCTCCA
104851 GTCCCACATT TCCCTTCCAC ACTGCCTTAG CAGAGGTTCT TCATGAGGTT
104901 CCTGCCCCTG CAGCAAACTT CTGCCTGGAC ATCCAGGCAT TTCCATACAT
104951 CCTCTGAAAT CTAGGTGGAG GTTCCCAACC TCAATTCTTG ATTTCTGTGC
105001 ACCCGCAGAC TCAACACCAT GTGGAAGCTG CCAAGGCTTG GGGCTTGCAC
105051 CCTCTGAAGC CATGGCCTGA GCTGTACCTT AGCCCCTTTT AGCCATGGCT
105101 GGTGCAGCTA GGATACAGTA CACTAAGTCC CTAAGCTGCA CACAGCAGGG
105151 GGGCCACACA GCAGGGGGAC TCTGGGCCTG GCCCATGAAA TCATTTTTTC
105201 CTTCTGGGCC TCCAGGCCTA TGATGGGAAG GGCTACGTTG AAGGTCTCTG
105251 ACATGCCTTG GGGACGTTTT CACCATTGTC TTGGCAATTA ACATTGAGCT
105301 TCTTGTCACT TATGCAAATT CTGCAAATGT CTGCAGCTGG CTTGAACTTC
105351 TCCCCAGAAA ATGGGTTTTT CTTTTTCTTC TTCCTTTTTT ATTATTATAC
```

FIGURE 3-31

105401 TTTAAGTTAT GGAATACATG TGCAGAACGT GCAGGTTTGT TACATAGGCA
105451 TACATATGCC ATGGTAGTTT GCTGTATCCA TCAACCTGTC ATCTACATTA
105501 GGTATTTCTC CTAATGCTAT CCCTCCTGTA GCCCCCGACT CCCCTGACCA
105551 GCCCTGGAAT GTGATGTTCC CCTCCCTGTG TCCATGTGTT CTGATTGTTC
105601 AATTCCCACT TATGAGTGAG AGCATGTGGT GTTTGGTTTT CTAATCCTGT
105651 TTTAGTTTCC TGAGAATGAT GGTTTCCAGC TTCATCCATG TCCCTGCAAA
105701 GGACATGAAC TCATCCTTTT TTATGGCTGC ATAGTATTCC ATGGTGTATA
105751 TGTGCCACTT TTTCTTTATC CAGTCTATCA TTGATGGGCA TTTGGGTTGG
105801 TTCCAAGTCT TTGCTATTGT GAATAGTGCT GCAATAAACA TATGTGTGCA
105851 TGTGTCTTTA TAATAGAATG ATTTATAATC CTTTGGGTAT ATAACCAGTA
105901 ATGGGATTGC TGGGTCAAAT GGTATTTCTG GTTCAAGTTC CTTGAGGAAT
105951 TGCCACACTA TCTTCCACAA TGGTTGAACT AATTTACACT CCCACCAACA
106001 GTGTAAAAGC GTTTTTATTT CTCCACATCC TCTCCAGCAT CTGTTGTTTC
106051 CTGACATTTT AATGATTGCC ATTCTAACTG GCGTGAGATA GTATCTCATT
106101 GTGGTTTTGA TTTGCATTTC TCTAATGACC AGTGATGATC TTTTTTTTTA
106151 TATATATTTG TTGGCTGCAA AAATGTCTTC TTTTGAAAAG TGTTCATATC
106201 CTTCACCTAC TTTTTGATGG GGTTGTTTGT TTTTTTCTTG TAAATTTGTT
106251 AAAGTTCCTT GTAGATTCTG GATATTAACC CTTTGTCAGA TGGATAGATT
106301 GCAAAAATTT TCTCCCATTC TGTAGGTTGC CTGTTCACTC TGATGGTAGT
106351 TCCTTTTGCT GTGCAGAAGC GCTTTAGTTT AATTAGATCC AATTTGTCAA
106401 TTTTGGCTTT TGTTGCCATT GCTTTTGGTG TTGTAGTCAT GAAGTCTTTG
106451 CCCATGCCTA TGTCCTGAAT GGTATTGCCT AGGTTTTCTT CTAAGGTTTT
106501 TATGGTTCTG TTGCATCGTC AGGCTGCAAA TTTTCCAAAC TTTTATGCTC
106551 TGCTTCCTCT TGAACACTTT GCTGCTTAGA AATTTTTTCC ACCGGATACC
106601 CCAAATCATC CTCTCAAGTT CAAAGTTCCA CAGATCTCTA GGGCAGGGAC
106651 AAAATGCCAC CAGTCTCTTT GTATAGCAAG TGTGACCTTT ACTCCAGTTC
106701 CCAACAAATT CCTCATCTCC ATCTGAGACC ACCTCAGAAG CCATTCAACA
106751 GGTCTCTAGG AAGTTCCAAA CTTTCCTACA TCTTCCTTTC TTCTGAGCCC
106801 TCCAAGTCTC TAGGAAGTTT CAAACTTTCC CACATTTTCC TATCTTCTTC
106851 TGAGCCCTCC AGACTGTTCC AACCTCTGCC TGTTACCCAG TTCCAAAGTT
106901 GCTTCCACAT TTTTGGGTAC CTTTACAGCA GCACCCCACT ACCCAGTACC
106951 AATTTACTGT ATTAGTCTGT TCTCATGCTG CTAATAAAGA CATACCCGAA
107001 ACTGGGCAAT TTATAAAGAA AAGAGGTTTA ATGGACTCAC AGTTCCATAT
107051 GGCTGGGGAG GCCTCACAAT CATGGTGGAA GGTGGAGGAG GAGCAAAACC
107101 ATGTCTTACA CGGCAGGAGG CAAGAGTGTG CATGCAGAGG AACTGCCCTT
107151 TATGAAACCA TCAGATCTCA TGAGACTTAT TCACTATCAT GAGAACAGCA
107201 CAGGAAAAAC CCACCCCCAT GATTCAATTA AGTCCCACCA GGTCCCTCCT
107251 AGGACACATG GGGATTATGG GAGTTACAAT TCAAGATGAG ATTTGGATGG
107301 GAACACAGCC AAGCCATATT ATATACTTTA TTAACTACTT ATCTCTTCTA
107351 ATAATACACT TGTTCTCATG TGGTTATTCT TCTTATTACC TTTTTTTTGA
107401 GATGAAGTCT CACTATGTTG CCTAGGCTAG TCTCAAACTC CTGGGCATGA
107451 GCCATCCTCC CACATGGCCT CCCAAAGTGC TGGGATTACA GATGTGAGCC
107501 ATCACTCCTG GTCTCATGTG GTTATTACAG TGAATGATTA AGACACCTTC
107551 TCTCCTTCAC TTTGTCCCTA CTTTTTCTCA GATTTTTTTC AATGATGTCT
107601 ATGCTTTTCT TGTTTTTTTC TAGCTTTTCT AACCTTGCAT TTATTTTCCT
107651 TCAGATCTCA ACATCTGCTG CAATTTGACT GTTCAGGTTA GTGACAATTT
107701 GCCCATTTAT CAGTTTTGTG CCTTAGGCAG TATTCACCAG CATTCTCCTA
107751 CTTGAGATGA ATAAGGATCT TTATTTATCT GACCACTTGT TTACTCATTC
107801 ATGGGGACAT TTAATATTTA CAGAACACTT TCATCAAAAC AAGCCTGTTT
107851 TTTCTTTTCA AAATATAATA TACTAGCATA GGAACTTGAC AGAAGAGGTA
107901 ATAATACAGA AGAAATCTAG AGAACTGATC ATGGAGAAAT AATTAAACTA
107951 AAACAAAGCT GCTGCTTATA GTAAGGTAGA CCAAGTTTGT CCTGTGTTCC
108001 AAATTATACT TAGCCAAAAA TAAATATTTA TAGATAATTG AATAGTAGTT
108051 TTTAGAAATG ATTCATGGAT TACTCAGGGG TGGAAATTAT CCCTGTAATG
108101 TAGGCCCCAA ACTTCTAAAA TATTTATAAT TTGTGAGGGA GAAATAAATC
108151 CACAAACATT TGAAAATATC TAATTTTAAC TTAAATTGAA AAACAGCAGT
108201 AGTGTATTTT TGTTAGGCCC TTATATTGGG TAATATAGTC TATGTAAGTA
108251 TGGGAAAGGC TGGTGAAAAC TGTGGATTTA TCTACAAAAT ACATTGGTAC
108301 ATTGGGGAGT TTTGTGTGGG AAGTGTCCTA ACACCTCAGG AAATGCTAAG
108351 GAAGAAATGG GTAGGGTAGA CTGAACACTC CACAAGGGTA GGGAGAGCGT
108401 CTTCTCTTCA TTGCTCTAGC CCCAGCACCT AGAATGTGCC TGGCACTCAA
108451 TCTTTTCTTT CCTTTTTTAA CTGATTTCAG GTTGAGTTTA TAGGCTAGCT
108501 GAGTTTATAG GCTGGTAGAG TACAAGGTCT ACTAATTTCT CTACTCATGA
108551 CTACTTTACT TTCTAATAGC ACTTTATGTT CTTCAAGTGC TTTTTTTCTT
108601 TGTTCTTCAA TGTATAATAA GGTGAGAACA GTTTTATCTC CATCTTGCAG
108651 GGAGAAGAGT TAGTTAAGAG ACTTTCTCAG GTCACACACA TATAGTTAAT
108701 GACAAGGTGA GGTTTAAACC TTAAATAATA GAAATAAAAG TGATTTTATA
108751 ATTATCTAGA GTAGTTTCAA TGTGAAATAA CTTAAAGGTA TGGAAATGGA

FIGURE 3-32

```
108801 TGCCAAGAAG TATAGTCAGT CTTGCTGGAG TAAAATAATG CCCAGTGCTT
108851 TGTGCCTTCT CCCAGCTGCT GCTTCCAGAA GAACGGGGTG TCTGAGTGTG
108901 AACATCACCC AACAAGTAGG TTAACAGATA TCCACGCCCC TCTTGACCCA
108951 CATACATATC AGTGGGATTT AGAATGCTGC CACATATTGA TGATTGAATT
109001 TATGAAGCAT ATAATATCCT CAATAATAAA CCAAGTGTCC CTGTCCCAAC
109051 TTGTTATCTC TGCTTCTGTG AACACATGTT TTCTTTTATA TGCTCCTTAC
109101 TCCTCAGGTG CTCTCTCAGG GACTTTTCAG TTCCTGACCT TGTCCTTTTC
109151 AGCATTTTCT CAGAGGACAA TTCTTAGCTT CCTGTTGATT CCTCAAGCAT
109201 TAATTGCTTT TTTCTGCCAG ATATTTCCTT GCTAGGCTCT TGAGCCCTCA
109251 GAGCTGTTCT GAATTATGCA GTGGGAATTG CCCAGGATTA GGAATCACCT
109301 AATGTCCCCA CCACCCCCGC TTCCTTGTGA GGCACTTCTC AACTCTGCAT
109351 CCCTTATACC TTCACAGCAA CCCTGTGTAC CCAAAGCAGT GTCATACTCG
109401 GTGCCTCCTT TTCTCTCCTG AAATAAAATT CCTAGATAAG AAGACCTCTA
109451 TATTCCAGGC CTTGTCTTTG ATTTTAGGGA AAAAAAAGAA AACTACCTAT
109501 ATACATAATG TTTTTAAAAA TCAGTAATGT CCCACTCGTT ACAGAAAGGA
109551 GAAATAAAAG AAGTAAGTTA ATGCCTGGGA TACGTGCTAC AACATGGATG
109601 AACCATGAGG ACATTACACC AAGTGAAATA CTCCAGGCAC AAAAGCACGA
109651 ATACTGTATG GTTCCGCTTA GATGAGGTAC CCAGAGAAGT CACATTCATA
109701 AATACTGAAA GTTGTATGGT GGTTTCCAAG GGGAGGGGGA AATGAGGAGT
109751 TATTTAATGG GCACAGAGTT TCAGCTTGAG AGGAGGTGGT GACAGTTGTA
109801 CAACAATGTA AATGTACTTA ATATAGTACA CTTAAAATGG TTAAAATGGC
109851 AAATTTTATG AAATAGGAAT TTATCACGAT AAAAAATTAA AAAGTAAGAA
109901 AAGTTACTGC TTGGGCGAAA GTATATCAAA AAAATAAAAA TAGTCCCCAC
109951 AAATTTCCAA AACAACCCTA ATGAGGTGTT GCTGCCTAAA TGGTGAACCA
110001 AATTGTGAAC CAATGTGTAG TGTTTGAGAC TGGGAAACTG ATGCCCAAGA
110051 TTTTAGCCTC AATAAGGAGT AGAGTTCATA ATTTGACTCC AAAGACATTT
110101 CTTTCCCTAC CATGCCAAGG CCATCTGATT CCCAGTCCAA AGAAGTTTTC
110151 TCTCTGCTCT GTAGGCTGCC TTAATCCAGA GTACACAAGC CTTCCATTTT
110201 CTTATCTGTC CTCCTACCAG GGTGTGGTCC TTTTCCTCCT GAACACTGAC
110251 TGTATAATTA CCAGACAAAA CTAAACATAT TTAAAATATA GGCAGTCCTC
110301 TACATCCAAG GTTCCACATC CTTGGATTCA ACCAACCATG GATTGAAAAT
110351 ATTTGGGGGA AAAAAAAACA ATAAAAAAAC ACTGGCCTGG GCAGCATAGT
110401 GAGATGCCAT CTCTACAAAA ACATTAAAAT ATTAGCTGAG CATTCCAGCA
110451 CTTTGGGAGG CTGAGGCAGG CAGATCACCT GAGGTCAGGA GTTCGAGACC
110501 AGACTGGCCA ACATGGCGAA ACCCTGTCTG TACTAAAAAT ACAAAAATTA
110551 GCCAGGCATG ATGGCAGCTG CCTATAGTCC CAGCTACTCA GGAGGCTGAG
110601 GCAGGGAAAA TTTCTTGAAC CCGTGAAGCA GAGGTTGCAG TTAGCCAAGA
110651 TCCCACCACT GCACTGCAGC CTGGGTGACA GAGTGAAACT CTGTCTCAAA
110701 AAAATAAAAA TAAAAATAAA TAAAAATTAG CTGAGCATAG TGGCATGTGC
110751 CCATGGTCCC AGCTACTTAG GGGGTTGAGG TGGCAGTGAG CTGTGATCGT
110801 GCCACTGCAC TCCAGCCTAG GCAACAGCGA GACCCCATCT CAAAACAAAA
110851 ACAATAAAAC AGAACACAGA TTAAAAACAA AATACAGGCT GGGCTCACTG
110901 GCTTATGCCT GTAATCCCAG AACTTTGAGA GGCCAAGGTG GGAGGATTGC
110951 TTGTGCTCAG GAGTTTTAGA TCAGCCTGGG TAACACGGCA AGACCACATC
111001 TCTACAAACA ACAACAACAG GAGACTATAC TTTCAGGGAC CATTTCTGGG
111051 GATCATAGTT TGTTACTAGA GAAGTTTCTC TGTGTAGAGC ATTGAAATAT
111101 AAAAATGCAG AATAATCATT TACATAGCAT TTACATTGTA TTGGTTATTA
111151 TAAGTAATCT AGAGATTAAT TAAAGTATAC AGGAGGATAT ACATAGGTTA
111201 CATGCAAATA CTACACCATT TTATATAGGG GACTTGATCA TCCATAGATA
111251 CGGGTATCTG AGGAGGTGTT GGGTTCAGTT CTCCACGGAT ACCAAGAGAC
111301 TAATGTTAAT TTCATTTCCC CAACCTCCAC ACCAGAACTC TGAAATAAGA
111351 ATAAGAAAAG GAGCAGTTGG GATAGACAAT ATCAGAAGTA TGTGGAAATG
111401 ATAACAGTGG AAGGAAAGCT GATCTAGGCC TACTCAACAA ATTTTAATCT
111451 TCATTCTGGT AAAAACAAAT TAGATTTATG GGTGCAAATT TGAGCCAGCA
111501 ATTAGATGGC TCTTAGGATT AATAAAAAAA GACTGAACAT CATGCCTTCC
111551 AAAGACTGAG GGAAAGAGAT AGATAGGAGA CTTTGGCAAA GTAGCACTTT
111601 AGCCAACATC ATTAGCCTAA ATCTTAGTGA AGAGAGGTTA GAAGAAAGGT
111651 AGAATTTTCA TGGAAGGATC CATTTTTCTT CACTTCAGAA TTAAGGGAAA
111701 AATTAGGAAG CTGAATAAGA ACTAATGGCC TAATTTCTTT GTTTCTTTCA
111751 AAAATCAAAT CTTTAAGTTA AAATTTCAAT AACCAACAGA AGAAGGTAGA
111801 ACCAATTTTG TGAATTTACA AAATACTTGC TTATTGACAC TATTGCCAAG
111851 GTCATTAAAT AAGTTATCAA GTTACCTCAG ATCCCAGTGA TTTAAATAAT
111901 GCTTTCTGAA TGTATCCTTT TCTGTTTTAA GAAGAAGCTG TATTAGGTTC
111951 TTGTGTTCCT ATAAAGAAAT ACCTGAGGCC GGATGATTTA TAAAGAAAAG
112001 AGGTTTAATT GGCTCACGAT TCTGCAGGCT GTATAGGAAG CATGGCCCCA
112051 GCATCTGCTC AGTTTTCTGG TGAGGTCTCG GGGAGCTTTT ACTCATGGCG
112101 AAAGCAGAGT GGAAGCAGCA GGTCACTTGA TGAAATTGAG AGCAAGAGTA
112151 TGGGTGGGGA GCTGCCATAC TCTTAACCCA ATCTCTAGTG AACACAAGCA
```

FIGURE 3-33

```
112201 ATAACTCACT TATCACCAAG GGAATGGTGC TAAGCCACTT GTGATGGATC
112251 CACCTCCAAA ATCCAGTCAC CTCCCACCAG GTCCCACCTC CAACATTGGG
112301 AATCACATTT CAACATGAGA TATGGAAGGG ACAAACATTC AAACCATATC
112351 AGAAGCCTAT CTTAGGCTGG GCACGGTGGC TCACGCCTGT AATCGCAGCA
112401 CTTTGAGAGG CCGAGGCAGG CAGATCATTT GAGGTCAGGA GTTTGAGACC
112451 AGCACGGGCA ACATGGTGAA ACCCCATCTC TACTAAAAAT ACAAAAACTA
112501 GCTGGGCATG GTGGCACACA CCTGTAATCT CAGCTACTCG GGAGGCTGAG
112551 GCAGGAAGCT CTCTTGAACC CGTGGGCAGA GGTTGCAGTG AACTGAGATT
112601 CTGCCACTGC ACTCCAGTCT GGGCAACCGA GTGAGGCTCT GTCTAAAAAA
112651 AAGAGAAGCC TATATTAAAC TTATAAAATT TAATATCATT TCAACTAGCC
112701 TTTTGTTGGG TGCATTTGTT CACTTTGGAC TATTTTTCCA AATTCATGTA
112751 CAGTCGTGCA TCTCTTAACA ATGAGGATAT GTTCTGAGAA ATGCATCCTT
112801 AGGCAATGTC ATTGTTGTGC AAACATCATA GAGTGTACTT AGACAACCCT
112851 ACATGGTGTA GTCACTACAT ACCTAGGCTA TATGGCATAG GTAGAGCCTA
112901 TTGCTCCTAG GCTACAAACC TGTACAGCAT GTTACTGCAC TGAATGCTGT
112951 AGGCAGTTGT AACACATGGT ATTTGTGTAT CTAAACATGG AAAAGGTACA
113001 GTGAAAGTAC AGGATTATAA TCTTATGGAA TCGCTGTTAT ATATGTGGCT
113051 CATCTTTGAC CAGAAATGTT ATTATGTAGC ACATGACTGT AATTCCACTA
113101 TTGATTAGAG ACTCCATAAC CCACACCTAC CTGTTCATTT GCATGCACAT
113151 TTAGCCGATA GGTGACTTTA TCACTAGATC AGTCCAAGAA TAAGTTTAAA
113201 GAGCATGTCC CCTTTTCTTG TTTTTTAATA AATGTAAAAC ATGGAAGTAC
113251 ATGATACCTC AAAGGAGCAT GAATCACAAA CAGGAGTCCA CATTTACAGA
113301 GCCTTTGCAA CCACTTTTAT ATCTTCAAAG TGCTCTACCA AAGTATCGGA
113351 GTCAGTCAGT CATGCAGGTA GGCTGCCAAC CATGGCTGAG CCCTCAGGAG
113401 CTGTCTTGAT GACCAGCATA GACACGGTCT GTTGTGTTCA TTCAAACTTG
113451 CCAGGAACCA AGGATCTGAG AATGGTAAGT CTGGTGTGCC AAGAATGAAA
113501 CTCCAAGAAA GAAACTTCAG AATCAAATTT AAATTTAGGC TGGGTGTGGG
113551 GGCTCATGCC AGTAGTCCCA GCGCTTTGGG AGGCTGAGGT GGGCAGATCA
113601 CCTGAGGTCA GGAGTTTGAG ACCAGCCTGG GCAACATGGT AAAACTCTGT
113651 CTCTGCTAAA AACACAAAAT TAGCTGGGCA TGGTGGTGGA CACCTGTAAT
113701 CCCAGCTACT TGGGAGGCTG ACGCAGGAGA ATTGCTTGAA TCCGGGAGGT
113751 GGAGGTGGCA GTGAGCCGAG ATCACACCAT TGCATTCCAG CCTGGGTGAC
113801 AAGACCAAAA TGCCATCTCA AAAAAAAAAA AAAATCGAAT GTAGGTGGAA
113851 TTAATAAAAA TTTAAATAAA CTAACAGATT AAACCTTTCA TTGTTCAGGA
113901 AGGAACAGAA CAATTTTTGA TAACTTGTGA AATATCTGGC ACAGAAATTA
113951 TTTAGAGCCA CTAAATAATT TCAAATTACC TAAAAATCCT AGTGATTTAT
114001 TTTCTATTTT AAGATGAAGT CTACTTTAAA CTTCTAAAAT GCAGGGTTAT
114051 TTAAACTGGC ATCTAAATCC AAGCTGGTTT TGGTTGGTAA TTCCTCTAGG
114101 ACATTTTACT AAATCTTGAT CTTATCTAAA TGATGCTATG TCATAGATGG
114151 ACTGTTTGTT TTGTTTGTAA TCCAGGGAAA TTAAAAAAAA AAAAAAACAA
114201 GTAGAAATAA AGGCTTTTAA AGAATTTTTA GAGTTAGAAA TGTTTTCAAA
114251 ATTAGGTTCT TTAAACCATT AGCCATCTCT TCCTTCTGAA CTCTTCTTTT
114301 TTCTGCCCTT TGGTAGCTAT GAAATAATCT GCATTCCAGA AACTTCTTTT
114351 TTCCCAGTCC TTTTTCATGT CTTAACAGTG CCATGCATGA TTATCTACAC
114401 CATGGAAACC CATCTTAATG AAATGGAAAG ATCTGCTGTT TAAAAAAACA
114451 AACAACCAAT CACCCATGCC TCCTACAGTC CTGTTAGGCC AAAGGTCTGC
114501 TCTCCACCTG GTCAGCTGCT GAGGAGGGGG CAGGTACCTG TACCACAGTT
114551 TAGCCACAAC AGAAATATCG TTGGGCTACC AAAAGATGTC AGTCCTTAAA
114601 AATGTTGGCA AGCAAAGAAA TTTTTCTCTA AGCACATCTT AGAGATACTT
114651 TAGTCTAAAG GCACAGTTGA TCAGATTAAA GAAGCTAAGA TTACATTTTG
114701 GGGCTGGTCT AATCGATAAA TATCTATCAG TGGCCATAAA GTTTTATTTC
114751 AGTGTTTTCT GGTGTAAAGG GATTAATTTT ATGTACTCAA GTCTATGGAG
114801 TTTGTTTTCT TCCACCTTTA ATCACCTCAC TCGGCTTGCT GGTGTGCAGT
114851 GCTCATGCAA AACTTCCTAA GGGAGTGCCC CACTTGTTTT AATTTACTTC
114901 CTCTAGCAAA AACACAGCTA GAGGAAACAC AAATCGTTTA AAAAAGAAAG
114951 AAAAAAACAA AATAAACTGT CTCATTTCAT TCAGAAGATT TCTGTTCATC
115001 CTACTTTTAA GTCAAAGGGT AGTGTTAAGT TTTGATTTGC AACAAACTCA
115051 GATTATCCAT TAACCCACGC TTGCAGGCTT TCTACTCCCT TTCAGATGTA
115101 CGTCTATTTT TTTAAAGGGC CCAGGCTTTT TGAGTTTACC TTTCAAAAAA
115151 GTATGTTTAG GCAGAATGGT CTTCTAGAGT TATGGAATAG CTTTTCAAAT
115201 GTTGGATGGA TATTTTCTTA GAATTACTCC AAAATGCAGT AAATGAGAAG
115251 CAACAAGCTG AAGCTGCCAG TCTCTGGTAT GAATAGCAGT GAGTAAAAGT
115301 GCTTTGAACT TAGTGTCCGT ATTGGAAAGC AGGAAGTTGA AAGAATCAGT
115351 TTGAATTCCA TAATTGCTGT AAATATTATG TAGCATTTAT TATCCCATTT
115401 TTTGATGTAC CATGTAAATG TAAAACATAC AGGTCCCCTC CCTCCCCTGA
115451 AGAGTTTATT CTCCTAGACA TAAAATGAGT GTACATGCCT AGCCATTGTC
115501 CACCCCTTTC CAGTTCGATG ATGTGTATCT GATGAGGAAG GACAGAGATG
115551 AGAGAATCAA AATCATTATA AACGAATACA GATGATGACT GGGTGACTGA
```

FIGURE 3-34

115601 AATCTTATCT CCCAAAGAGC ATAGTTAACT GCAGCAGCAG TGGATGATAA
115651 CTACTGGGTG GGGGTAGAGG GGTTGTTTTG CTCAGTTCTG CCTAGAAGAC
115701 AGTTGTAGTT ATTTTAGTCC CACAGTCTCT ACTCCTCCCT GGGCCTGTTT
115751 TTGCCCTGCT TTCCTGGCCT CTGTATACGG CCTTTGTTAG CACTTTGTAG
115801 TTGTCAGTCA GCTGTTACTC CAAGTCCTCT GTCAGAAATA AAGTTACTTT
115851 TGTTTGAAGA GTCTCCAGTA ATCCCCCTTC TTTTTAAACT ATGACTCCCC
115901 AAAGATATAT AGTCTAACTT GTGTGCACCA GTATTTTATC ATATTTATTT
115951 AATAAATTCC ATAAACCTTA CATAAAAGAT ATTAAACAAA AAGTACTCAC
116001 CCCCTTAAAA GGAGGAAGAT AATGACCATC AAGGCATGGC CAAATTACAC
116051 AGTCTTAGGA AACAGTACAG TAGGTAATTA TTGGACCTTT TTGATTAATG
116101 GCATCCTTGC TTTGCAGGTG TGTACCACCA CTGCACTTCA CAGGTTTCTG
116151 TAAATAAAGT GATTTGGGGT TCCCTTCTGT TGAGCATTCT TCTATCGTGA
116201 AACCATTTGG TGATGAGAAG CTTGAGTTTC TAATGCATGT TCGTTGGCTT
116251 TATTGGAGCT GCTTTTGTAC TGTGGCCACC CATTTAACTT CCTCTGTGGT
116301 TGATATATGA GAGCTAGAGG ACTCTGCCTT AGCTTCTTCA TCAGAACACC
116351 ATGTGATGTC TCCAAATGTC TTACACTGTG GCTCCCCTGT AAGAGAGGTG
116401 GAGAGGAGAG AGGCCAAGGC TGAGGCCATT CCAGTGAGAT GATCACTTTC
116451 TTAGGCTACT GCCTGAGACC TCTCAAACAG GATCTGTCTG TGGGCTCTAG
116501 TGAGGGGAAA GTAACCGGCC TTCAGTTGGC TGGTCAGAAA CAGTTACCCA
116551 CGGCCGGGTG TGGTGGCTCA TGCCTGTAAT TCCAGCACTT TGGGAGGTCA
116601 AGGGTGGAGG ATTGCTTAAG CCCAGCAGTT TGAGACCAGC CTGGGCAACA
116651 TAACAGGACC ATGTCACTAC AAAAATGATT TAAAAATTAG CTGGATGTGG
116701 TGGCATGTAC CTGTGGTCCC AGGTACTCGG GAGGCTGAGG TAGGAGGATT
116751 GTTTAAGCCC AGGAGGTCGA GGCTGCAGTG AGCTGTGATT GCGCCAGTGC
116801 ACTCAAGCCT GGGCAACAGA GTGAAACTCT GTCTCAAGAA GAAAAAATGT
116851 TTCAGGCACA GTCAGTGTTG AAGATGTCAA TTAGCAAGGT TTTTTTAATC
116901 CCCTGGAAGT GTGCAACCCA GGATTAAACA GATGTTTAAG AACATAAATA
116951 AAAGATGAAT TCCTGGCCGA GCGCGGTGGC TCACACCTGT AATCCCAGCA
117001 CCTTGGGAGG CTGAGGCAAG TGGATCATGA GGTCAGGAGT TCGAGACCAG
117051 CCTGACCAAC ATGGTGAAAC CCCGTCTCTA CTAAAGATAC AAAAATAAGC
117101 CGGGCGTGGT GGCGGGCACC TGTAATTCCA GCTACTCAGG AGGCTGAGGC
117151 AGGGGAATTG CTTGAACCCG GGAGGCGGAG GTTGCAGTGA GCTGAGATCA
117201 TGCCACTGCA CTCCAGCCTG GGTGACAGAG TGAGACTCCA TCAAAAAAAA
117251 AAAAAAAAAA AGATGAATTC CTGACACTTA ATACATTTAG ATAAAGTCTA
117301 ACTTATCTTT AAAGTTAACT TTAAACAACT GCTCTACAGC ATAATATTAT
117351 CCCCTGTGTA GACAGTGACC TCTAGTATAA GCTTTAAATT GGTTTAAAAA
117401 ATAGATTAAG TATGAAAACA TCTCTTCAAC ATTTTCATAT CTCTACAACT
117451 CTTTATTAGA AATATCCATT TTGTCCTATA TTCACTTTCT GCTGAACTAA
117501 AGTGCATGCC AGTCTCCAGA GGATGACAGA TGGGCAGAAA TTTATAAAAC
117551 AGCAGCAGTC ATGGACAGCC TAGGCCACAA GTAGAACATA CCAGACCTCC
117601 TGGGTACTAC TGTCTCAGTG AGAAAGGATC CAGAGATCCA GCATCACCAG
117651 ATTCCCTCAC ATACTGACCA AAGAAACTTT TTTAGTTTGC AGATTTTGGT
117701 GGGAATAGGG AGGAGTATTA GAAGCTGAAG TTCTTGCTTT ATTTAAATAA
117751 TCATGGCAGT AACAAAATTC GGTTAAATAA TGCCCCTTCT AGGCCAGGCG
117801 TGGTGGCTCA TGCACTTTAG GAAGCCGAAG CAGGCAGATC ACTTGAGGTC
117851 AGGAGTTTTG AGACCAGTCT GGCCAACTTG GTGAAACTTC ATCTCTACTA
117901 AAAATACAAA AATTAGCTGG GCGTGGTGGT GCACACCTGT AATCCCAGCC
117951 ACTCTGGAGG CTGAGGGAGG AGAATCGCTT GAACCAGGTA GGTGGAGGTT
118001 GCAGTGAGCT GAGATCACAC CACTGCACTC CAGCCTGGGC GACAGCGAAA
118051 CCCCATCTCA AATAAATAAA TAAATAAATA AATAAATAAA TAAAATAATG
118101 CCCCTTCTAG TTGAAAAACT AAGTTCCTAC CTAAGCATAA TTTGGATTTA
118151 CCCAATTTAT CTTCTTTCAA AATACCTCAA ACATTTCACC TTATTATTCT
118201 TTTTAAGGAT TACAAAGTAG AGCAGGGGGG AAATAATAAA CCACTAATAA
118251 AGAATAATAG CCATTTGACA GACAGGTGTT CTTAGTTTCA TAAAAAAAAA
118301 GATGCCTGGT AGATTCAGTC TTTTATGAAT ACTAAAGAAT GCCTCTTATT
118351 TTTGTTTGTT GGAGACAGGT TCATTTTGAA CCTAACCTGG TGTTCTCAGG
118401 GAACATAGGT TAGAGGGGAG AGATTTAGGA GTGAGGGCTC AAGCAAGAAG
118451 CATGTTAGAA GACTGCTGTA GTGGTCCCGG TGAGGAGTGA AAGGAATGGA
118501 ACTAAAATAC CATGAGGAAC GTGGGCCAAA GGAAACAAGT CTGAGAGATT
118551 TAGAAAGTAA ATCATCAGGA TTCAGTGGTG ACTCTTGGAT TTGAAGAGGG
118601 AAAGGGAATA ATCTAGGGTA GCTGTCAGTT TCTGCCATGG GTAGTTGGGC
118651 TAACAGTAGT GTATTAACTG AAATGGGGGG CAGGCAATTT GTGAGGGTGA
118701 GTTGAGTTCA GCACAGGGCC TGTTGACTTT GAGGTGCCTT TGAAACAGGA
118751 GTGGATATGT CTAATTTTAC ATGAAGTGTC TACTTAAGAG AATGTGTAGA
118801 GACATTAACA GGGCTGGGTA GGAACACGGA ATACCTCAGG TGCCAAATGA
118851 AAACTTTTGG CAATAACAAG AAGTTAGGAA GTTGAAGAGG GTAGAAGAGG
118901 CAGCAGAAAT GTAGATGGAT AAGAACATGT TGGCATGTTT AAGTGACATT
118951 GTGAATAAAC CTGTCCTACT GGGCCAGATG GGATTAGGGC TGAGATCAAA

FIGURE 3-35

```
119001 GCAGGCCTAA TGCAGTTTGC AGACGTGTTT TGTTAGACTT TTGTAGAGCT
119051 GGATCACACA GTGTCCTAAA TTTAAATTAG GTGCCAACAT TTCCACACAA
119101 AATCCGGATT TCTGACTTTT CTTTTAAACT AAAGGGCTCC TAGAGGTAGA
119151 TTTGGCAACA TTGGTAGACC TATATGATAA TAATCGACTG AAGTATATGT
119201 CCTTTCTCTC CAATGAACCT GTTTTACTTA TGTTATTTCG CTGAGCCCTT
119251 AGACATTTAA ATTTTGTACT TTTTTTTTTT AATCCAGGCT TATAAGTCAG
119301 ATGAATTTTC TACTTCTGAG TCAAAGATCA GTAGGTAATA AAGGTACAAA
119351 GATAGATTAG CAACAGATTA CGGAGAGCTT TGAATACCAA TCTAAGGAGT
119401 GTAAATGTAG GCAGTGGGCC ACCTTTGAAT AAGGAATTGA TGAGATTAAA
119451 GCCATATTTA GGAGGATTAT TCTGGACCAG TATGAAAACA CAGAAGTTAG
119501 GGAAAACAGT TAATAGTTTT GAAAGAGAAG AGAAAAAGGA GATGGTGTTG
119551 GGATACATAA ATGGGCTTTT AAAATGCAAA ATGAGAAGTG TTTTAAAGAG
119601 ATATCACCCA GAAAGTCTAT GCACTGCCAC ATGGGCACTA TATGGGTGGT
119651 TGTATTTGGT GGGAAATTTG CTTGCAGACT TCCAGAACTC AGACCAATGT
119701 GTGGTGTGGG GGACGGTGAT TGTCAGGCAT TATGGAAAGG TCAAACAAAA
119751 TATGCTCACT GGCTATCTAT GGCCCACAGG TCACTGTAGT CTCTGTTATA
119801 AGTACACTAA GTGGAGGAGA AAGGTCCTTT AAAAAAAAGA AAGCTAAAAT
119851 TAATACCTGA TTGTTATTAA CTGTGTGCCA AACACTGTTC TAAGCTCTTT
119901 ACACAGACAT TTTATTTAAT CCTCGCAACC AATTTCTGAA GTAGGTACTT
119951 TTCCAATTTC CATTTTACAG ACAAAGAAAC TGAAACCCTA GAGGTTAAGA
120001 AGTTATCCAA AGCCACAAGG CTGATAAGAA CAGAACCAGG ACTTGAACGC
120051 AAGCAGTCTG CCTCTCCAGA GGTTTATCTT TTAACTGCTA TGTTAAACTG
120101 CCCCTGCATT TTAATCTGTT CTAATGCTAC ACAGATAGGC AACTTTACAG
120151 GTAGAGGACC TTATGCTTTA TTCTGGATGC TCTGTTATAA CTCGTTTCAG
120201 GGGTGTCAAT TTGGTCCAGG TCCTCCTGGA AGAAATAAAA CTCCAGAATT
120251 GACCCTTGAA CTGTCTCTTA GGGAGCAGAT AATGTAACGG GTCCTTGGGT
120301 GACCTTGAGA GAACAGGTAT GTTCAAATGT CTGTTCTCTT CCTTTAGCTA
120351 ATGGATCAGT GTAGCTTATA ATTGCATGCT TCTAACCCTT TGTTGAAAAA
120401 TAAAAACTCT TATAAACATG CTTTTTTTTT TTTTTGGAGA CGGAGTCTTG
120451 CTCTGTCGCC CAGGCTGGAG TGCAGTTGTG CAATCAGCTC ACTGCAACCT
120501 CTGCCTCCCG GGTTCAAGCT ATTCTCCTGC CTCAGCCTCC CGAGTAGCTG
120551 GGATTACAGG CATGCGTCAC CACGCCTGGC TAATTTTTAT ATTTTTAGTA
120601 GAGATGAGGT TTCACCACGT TGGCCAGGCT GGTCTCGAAC TCCTGGCCTC
120651 AGGTGATCCA CCCACCTCGA CCTCCCAAAA TGCTGGGATT ACAGGCGTGA
120701 GCCACCATGC CCGGCCTTAA AAATGCTTTT AAAAATGAAA ACTAAAACAT
120751 GTTAATTTTT TTCAAATGTT TTCATGAAAA TTATCACAGG ACAAGTTTCA
120801 TAAATATTGA AATTTGGAAA AAGTTGCAAG CCTATAACAT TGCAGAGAAG
120851 CAAATGCATT TGATGCAAAG CCTCAAATTT GTCAAGTTTT TCTACCATAT
120901 TCAGTGTGGT TTCTTTCTCT TTGGCCTATA GATGAAACAT GTAAATGAAA
120951 GATTTCAAGA TGAAAAAAAT AAAGAGGTTG TTCTCATGTG CATTGGCGTC
121001 ACTTCAGGAG TTGGACGACT GCTCTTTGGC CGGATTGCAG ATTATGTGCC
121051 TGGTGTGAAG AAGGTTTATC TACAGGTACT TTTTTACACC TTTTTTCCCC
121101 TATCAAAAAT TACTCTCATC ACCCAATGTC TCATTAAATG TACTTACATG
121151 CTTAAATTCT TTTTTTTTCT TTCCTTTTTC TTTTTGAGAT GGAGTTTCGC
121201 TCTTATTGCC CAGGCTGGAG TGCAATGGCA CGATCTCAGC TCTCCGCAAC
121251 CTCCCACCTCC CGGGTTCAAG GGATTCTCCT GCCTTAGCCT CCCAGGTAGC
121301 TGGGATTACA GGCGTGTGCC ACCACACCAG GCTAATTTTT GTATTTTTTT
121351 AGTAGAGATG GGGTTTCTCC ATGTTGGTCA GGCTGGTCTT GAACTCCTGA
121401 CCTCAGGTGA TCTGCCCGCC TCGGCCTCCC AAAGTGGCTT AAATTCTTCT
121451 ATAAAAATGA GAAATATTTT CTACAACATA ACTTCTATAG GCAGTTTTTC
121501 AAGGACAAAA TTAGTTATTA GTTTGGGTTT TAAACATGAG AAATTGGCAA
121551 TGAAACAACA ATTTCTTTGT TTTGTCGTGG AACTCCACCA AACCAGAATG
121601 GTTTTCATCC ATTGCTTTTT CTATGAAGAA TGTTTTTTGG TGTAGTTCTC
121651 ATAGTCATGT GCAGATCCTG TGCCCTTTGC ATGTCTTATG AAATTTGGTT
121701 GTGTGTGTGA CTTTTCAGCT TCTTTACTGC AAATTGCCTC CTGGTGTTTT
121751 GGGGTGAGCA TAAACAAATG CTAATTCCAA GATCATTGCT GACAATCAAC
121801 AGAACAGGTA TTGAAGTGAC TCCTTCATGC CACACACTCT GCTAAATGCT
121851 GAAGACTTAA GTGAAACATG GTCTGTGCCC TCACCTAGTA GCTAATGGTC
121901 TCATGGGAGA AGATAAAGCA GTGTTGCAGC ACAGTGGAAT AACGGGTTTG
121951 CTAGATGAGT GCCATAGCAA CACAGGCACC ATGCATCTGA GTCACTAGTG
122001 AAGGGCTGAG CAGAGTTAAG AGGTGAGGAC CAAGTGTGTA GATAAGGGAG
122051 ACAGAGAAGC ATGCTCCGAG CAGAGACAAA CACATGGGAA ACACCTCCCA
122101 CACTGTGGAG GTGTGACATG GGATGTTATA TCTGGGAAAC AAATGTTTGA
122151 ATGGAATAAA GGGAGAATAG TGGATGGATT GGTTGGGGGG ATGGACAGGA
122201 AGCAGCTTGG GAAGGGCGGT TATGTACATG AGCTCTATCC TGCAATCTAC
122251 TAGGAGCTAT TAAAGGATTT TAAGCCAGAG AGTGACATAA ATTGGGGTGG
122301 CGGGGGTTGG TGTTTTTTGG TTTTTTGAGG CAAGGTGTCA CTCCGTTTCC
122351 CAGGCTGGAG TACAGTGGTG CCATCACAGC TCACTGCAGC CTCAACATCC
```

FIGURE 3-36

```
122401 CGGGCTCAAG CAATCCTCCC ACCTCAAACT CCTGAGTAGA TGGGACTACA
122451 GGTGCGTGCC ACCATGCCTG GCTAATTTTT GTTTTTTTGT TTTTTTTCTG
122501 TAGTAACAGG GTTTCGCTGT GTTGCCCAGG CTGGTCTCGA ACTCCTGGAC
122551 TCCAACGATC CACCCATTTC AGCCTCCCAA AGTGCTGGGT TTACAGGTGT
122601 GAGCCATGCC CAGCCTGAGT TGTGTTTTAT AAAGATCAAT TTCGGCTGTG
122651 CGTGGTGGCT CATGCCTGTA ATCCCAGCAC TTTGGGAGGC TGAGGCAGGT
122701 AGATCACCTG AGGCCAAGAG TTTGAGATTA GACATGGTGA AACCCCGTCT
122751 CTACTAAAAA TACAAAAATT ATCCAGGCAT GGTGGCACAT TCCTGTAATC
122801 CCAGCTACTC AGCAGGCTGA AGCAGGAGAA TCACTTGAAC CTGAGAGGCA
122851 GAGGTTACAG TGAGCTGAGA TTGTGCCACT GCACTCCAGC CTGGGTGACA
122901 GAGCAAGACT CCATCTCAGA GAAAAAAAAA AAATCAATCT GGTAACTGTG
122951 AAGAGTGGAA TGGAGTGGAG GAGGGGACAA GAGTTGAGAT AGTCACAAAA
123001 GGCCACATAC TGAATGATTC CATTTATATG AAGTGTCCAG AATAGGAATC
123051 TCTGTCTCTT TCCATAGAGA AAGAAGGTAG ATTTGTGGTT GTAGGGGCTA
123101 GGGAAAGGAG GGAATGGGGA GTAACTGCTA CTGTGTTTCT TTGTGGGGTA
123151 ATGAAAATGT TATAAAATTA GATAGTGATG AAACTCACTG AATTGTGCAC
123201 TTTAAGTGGG TGAATTTCAT AGAATATGAA TTATATCTCA ATAAAGCTAT
123251 TTTTAAAAAG GCCAGGCATG GCAGCTCATG CCCATAATAT CAACACTTTA
123301 GGAGGCTGAA GCAGGAGGAT CACTTGAGGC CAGTTCGAGA CAATCCTGGA
123351 ACCATAGTGA GACCTCATCT CTACCAAAAA AAAAAAATTT TTTTAAGTTA
123401 GCCAGGCACG GTGGCACATG CCTATAGTCT TAGCTATGCA GGAGGCTGAG
123451 GTGGGAGGAC TTTGAGCCTA GGACTTTGAG ATTACAGGGG AGTATGGTCA
123501 TGCCACTGTA CTCCAGCCTA GGTGACAGAG TGACACCATG TCTCTAAAAT
123551 TTAGAGAAAA AGAGTAGAGG TCCAGGGACT AGTTGGAAAC TATATTAAAG
123601 TAGTCTAGGC ATCTAGGCAG GAAGAACCAA GGCAGAGATA CTGAGATCAA
123651 ATGGAATAAG CTAGGGACAT AGAGAATGAG AATTTGATGA AGCCACTTAG
123701 GCTGGATCTG AGATTTGTGG CTCAGGTAAT GGAGTAGATG ATTTTGCCAT
123751 CTTATCAAGA TCAATATGCA GTAGATCTAA GTATGATGAT GAATTGGGGT
123801 TTTTTAAATT AAAAAAAAAT TTTCACCCAA CCTACCTGCC CTTGCCCCAG
123851 TGATGAGTTT GTTTTTGCCA TGTTGTTTGT GATGCCTGTA GGAAAGCCAA
123901 GTAGTGTTTA GAGCTCATGG AAGAGTCTGG ATAAGAGGCC TGTGGGTTAA
123951 GAGGCCTGTG GGTTAGGAAG CCTGTGGGTT TGGCATTCAT GGGCTGCATG
124001 AAGTAAGTGG AGGACTAAGG AGATGGAAAG GGAGTGGCCA GATAGGCAGG
124051 GGGAAAACAG GGATGAGAGT GTTTCTAGGG GCAGGAGTGG TCAGCAGTGT
124101 CAGTAGCTAG TGACAAAGAG GCTGAAAGGT TTTCATTGAA TTTAGAATAG
124151 GGAGGCTATG AGTGAGCTTA CAGAGAATGG TTTCTACAGG TGCAGAAAAT
124201 TGATCATGAT CATTTGAACA AATGGAAATA TAGACAACTT TGTCCAAATG
124251 CTTGGCTGTG GAAGGAAGGT AAAGGCAAGC CACAAGGGGG GGTCTTAGGT
124301 TTTGAGGATT AAGCCTGTCT AATTATATTA TACATGAGAG GCAGTATTTG
124351 AGCAGGGCCT TGAAAAAGAG GTGAAATTTG GACACAGGAA AATGATTGGA
124401 AAAGGCATTA CAGATAAAGT TAACTTCCAT TAATTGACTT GAAGTAATAA
124451 CAGTCAACCA TTTATTGAGG ACTTTCATGT GCCAGACAAT GAACTAAGGG
124501 CTCTACATGC ATTATCTCAT TGATCCTTGC CCCAGCCCTT TAAGAGAGAA
124551 GGTACCATTG TTATTTCCAC TTAGAGATGT GAACACTGAG GAACTGAGAG
124601 GCTACCTCAC TGTGGGTGTC TGTGTGTATA GAGGTCCAGG CAGTCACAGG
124651 ACAGTCTGGT CACACAGCTA GAAGGAAGTG AAGAGAGTTG GTAATGTGCT
124701 GCTTTTAAAA ATGTATTTAT TGTATCATGA TCACTTTGTC AGTACTTCAC
124751 TGTGGACATC CTCATCTAAC ATTTAGTTTT GTTCTCTAGT GTCAAGGGAG
124801 CCCTCTAATG GACATTTATT GCACTACAGA CTTTCAGCTT TCATACATTC
124851 AAAAATTGAG TGCCTCCTGT GCCCCAAGCA CCAGCTCAGA TGCTGTTAGG
124901 GTGATGCAAA CAAGACAGAC ATGGTCCCTG AGTTCCTAAA GCAAGCCTGA
124951 GGCAGGAAGA AGCCGAATGT GTGTGGAAAC CCAAGAAGAT GGGGAAAGTG
125001 GCATGGGAAG GGACTGGAAA GTTAGAGTGG GTCAGATTAG ACAGAGCCTT
125051 GAAAGCCAGG ATGAAGAGAC TTTGCTTTAA GAGTAGTGGA TTTTGGCCGG
125101 GCGCAGTGGT TCACGCCTGT AACCCCAGCA CTTTGAGAGG CCAAGGCTGG
125151 CGGATCACGA GGTCAGGAGA TCGAAACACA GTGAAACCCC ATCTCTACTA
125201 TAATTACAAA AAATTAGCCA GGCTTGGTGG CACGCGTCTG TAGTCCCAGC
125251 TACTGGGGAG GCTGAGGCAG GAGAGTCGCT TGAATCTGGG AGGCAGAGGT
125301 TGCAGTGAGC CGAGATCGCA CTACTGCACT CCAGCCTGGG TGACAGAGTG
125351 AGACTCCGTC TCAAAAAAAA AAAAAAAAGT AGTGGACCTT TGTTGGAGGA
125401 CGTAGGTCTG ATTTTGTTTT TAAAGTCACT GTGGTTGTTC TGTGTTTGGA
125451 GGATGGATTA CAGGGAGGGA AACAAGAGTG CAAGAGGGGA GGAAGACCTG
125501 TTAGGAGATC AATTCAGGGT CCAGGTGAGA GATGATGGAC TATGGTGCTG
125551 TAATAGATAA AGTCATGATA TATTTTATAT ATCAGCATTT TATTCCAAAA
125601 CAATGTTTGA ACGTCCTCCT ACCCTAGATA GGGCAGACAG ATTATCTGCA
125651 ACATTTTTGG AGCACATTTT AATACCTGAC TGTTTCCAGT AATTTACAAA
125701 AGAAAATATA GCCTTTCCTA ACTTGTCCCA TGTTGGTCTG CAGTTACACA
125751 GCAGTAAGTT AAAAGTTAGT ATTGGGGGTC AAATATTTCA CTTTAGATGA
```

FIGURE 3-37

```
125801 AAGTTTAGCC ACAATCTGGC TTCTGTTAGG CCTTATCTAA TTTTTGCATC
125851 CAAATGTAGA GCATCGTTTG TGGCACCCAG TAGCACATGC TGAGTCACAG
125901 GTGTGACAGC TGCATTTCAA ACAAGCCTGA GAAGGAGAAA GAAAGCCCTT
125951 CAGTGTGTCC TGTGGTTGCG AGGAGCCACT CACGGACTCC ACCTTGTGAA
126001 CACAGCGGCA CAGGACGCAA CACAGGCCTA ACCCATGCAG GATGCTGGAC
126051 TCGTTCCCTT ATTCACTACC TCCTCTCTCT CCTTTTTCAT GGCTTCCTTG
126101 CCCCAACATC CCAAACACAC AGTGTGGTTT TTGATTCTTG GCTCTTCTCT
126151 GCTGCAGTTG ACTCCAGCTC TGCCTGTTTG TTCCTTCCTT TTTCTCCATC
126201 CCTGGCTCTC TGCCTTTTGG CCCATCCTTT AAGGCTTGGA ATGCTCCTGG
126251 GCTTACCTTC TTTCCATTCA TTGGTGGTGT GTTTTCTCAG ACATACCTGC
126301 TCCCTGCTTT CATCTTTCAA CTTCTTGGGG CTGTGACAAC TCTTCCCTCT
126351 TTGTCCCCTG GAAGCCAGTT CTGAGTAGCA GCCAGGCCTA GAACACTGGT
126401 GACACAGACA CACTTCATAG CCCTCCCGCA TGGTCTAGTT TCAGATCATG
126451 GTAATCCCTA GTCTAGGAGG CTGCCGAGCC CAAGAGCACA GGCTCTGGAG
126501 TGAGAAGCCA GTTCACCCCA GTTCTACCAC CAACTTGAGA ATCAGCAGAG
126551 GGCTGTGGTG AGGATTCTTG GTGGCAGGTT ATGTAAAGTG CCTAGCCCAG
126601 ACTGGATGAT TAATAAATCC TTGCATCTGT TATGTTTTAA TATCTTATTA
126651 AATACTGAAA GCAGATCCTG ATTTGGAATA GGTCTCAAGA AAGGAGACTA
126701 GGGTCTAATC CTGAATTAGA GTCTTTGCTT ATAGGTAACA AAGAATTTAT
126751 GAATTTATCT CACATTTTTG CTTAAGAGTT CTGAATTTAA ACTTCCATCA
126801 AGGTCCTGGG TCCCAGGTGT TTTCAACCTA ATAATTCTAA ATATTGAGTG
126851 GTTGGTTGCA GTAGCTTATG CCTATAATCC CAGCGCTTTG GGAGGCCAAG
126901 GCAGAAGGAT CCCTTGACCC CAGGAATGCA AGACCAGCCT GGGCAACATA
126951 GTGAGACCTC ATCTCTACAA AAAATTAAAA AGTTAGCTGG ACATGGTGGC
127001 CTACACCTGT AGTCCCAGCT ACTTGGGAGA TTGAGATGGG AAGATTGCTT
127051 GAGCTGGGGA GGTTGAGGCT GCAGTGAGCT GTGATCATGC CACTGCATAC
127101 CAGTCTGGGT GGCAGAGTAA GACTTTGTCT CAAAAAAATA AAAAAATAAA
127151 AAAAAATGCT GAGTCAAGTC TACTGCTCCT GCCAGAAGAG ATGACTGAAG
127201 TGCATTACGT AGAATAATAA TGGTTAAAGA AAAGCTTTGC AAAAGTTCCC
127251 AGAATATATA CTTTATTGGG ACAGGAGAAG CTACGTGTGT CGTGGTATCT
127301 TTTTACTATT TTCTTAATCT TATAGGCCTG TGTTTCTAGT CACCATTAAA
127351 TTACTACAGA TTTGTGTTTT TAATGTAATA TATAAGTGTT TTGGAAGGGT
127401 GAGAATATTT CAAAGGTTTG AAAGTTAAAA CTGTGTATGA AAGAATTGAA
127451 AACTTGGAAT TTAGATCACT TTTCCATTGT GTCATATTTC TCTGTGACAT
127501 GGACATATTG AAGCATGGAC ATCATGTCCA TGCACATAAA GCAGACAACC
127551 CAGACAACAC ACACATGTGC ACAGGAACGC TCTGGAAGGA TGCAAACCGA
127601 ACTGTTAAGT GTTGATTCCT GAGGAGGTGA ATGGGCATGG GTTGAGAGTG
127651 AGAAGAGGTT GTAGGAAGAC TTTCATATAT TACTGTGTAC ATTTCTCTAA
127701 GGTTTGAATT TTAAAAAATA TATTCATGAG TTACTTTTGT AATGTAGAAA
127751 TATTAATGAC TTTCCTATCC ATCAGTCTGC CTAAGCTTCC TCTTCCGGTT
127801 CAGGTAGAAT GAATTGGATC AGTGTTGCTC CATTTTCCAT TTTAGCATTT
127851 TACATTTGCC TAAAGATATC TTGGGATGAG GGTATATACT TTATCAAATG
127901 TAAGCTATTT CCAAAGTAGT AAATCCAAAT ACTTAACAAC TTCCCAGCCT
127951 AAGTAAATGA TCAGAGGCTC CGTACTCAGT CTCATCTAGA CTGTGGCACT
128001 GGGTGTGAAC GTATCAAATG CATGTTTCTC CATCAGGCAG AAGTGAGAGT
128051 AACCATGTGC CATGGAGAAG GTTGACAGAC TCCCTGTGAA GCACTTCGAA
128101 GTGACACTGG CCTCTGTGTG CTTCAGAAGA ATCCAGCCAC CTGCTGTGTG
128151 GCCTGACATT TTCCTTTAGT TTGTGATGGG CCAGCAGAAC TCTGTTGCCA
128201 ACTGTTTTCT GTCCTGGGTG CCTAGCCAGA GGTTCTGAAA GTCTGGAGAC
128251 TTTATATTGG CTAAACTTTA GGAACGTCAA TTACATGTCT ATCTCCAAGA
128301 TGCCTTCTTT TATTCAGGTG CAGCTCATTG TTTCCTCTTG AGCTACACTT
128351 AAGATTCTTG AGCAAAACCT AAACTGACAT TTCTCCAGCA ATGCTCTCCT
128401 TGAGATAGAA ATGGGAAAAG TAAGAGCAAA AGGAATCTTT TGTTCTCATG
128451 TGCATACACT AACTCATAGA AGGTTAATAC TTCTATAGCC TGTACTATTA
128501 TAACAAGTAT TATATATTTA TGATATATTT CCTTAAAGAA AACAAAAGCA
128551 ATATAGACAT CTAAACTGTC ACTGGCTTAT TAAGTGTCAG TGCCAGAGCC
128601 TAGGAGAAAA TAAGGAGCCT GTGAATTCCT TACTCGAATC TAACCAGAGC
128651 TGCTGTGTTT GAGAGCAAGT TTTAAAAGAT TGTATGTAAT ACTAAGTTTA
128701 TTCATCTTTC ACACTGAGTC CCAGCATCAC CAGATCAGTA TTTGATGCCT
128751 GGATCAATCT TTATTCTGGG GAGTGATGAA GCATTGAACC TGCTATATGT
128801 ATAGTTTGCC GAGCGTCGGC ATGTGCTCCT TGTGGCCCAG GCATCCCTGC
128851 ATATAAGGAA TAGGTACGTT CTCACGAGCC TCACCTACTT ACCTCCACAT
128901 TTAGCCAGAT TCTGGGTATT AACATCTGCT GGGAAAGAGC ATCACTACAG
128951 TAGCTACAAA TAAGGTGGAA GAAGCAAAGT ATTTTTCTGA GAAGTACTTA
129001 AAGAATAGAT GTGTAAATTT CTATAAACAC AAGTCTTAAA GGAAAATGAA
129051 AAAATTTTAC ATTTAAATAA CTACATAAAT CATTGCCTAA TTTTAATAAG
129101 AATATAACTT AATATAGCTT GAATGGAGAA AAGGACAACT TGCAGTCAGG
129151 GAAAGTATTA AGAAATAATA TGCTCAGTCT GGGCGCGGTG GCTCATGCCT
```

FIGURE 3-38

129201 GTAATCCCAG CACTTTGGGA GGCCGAGGCA GGCAGATCAC GAGGTCAAGA
129251 GATCAAGACC ATCCTGGCTA ACACAGTGAA GCCCCATCTC TACTAAAAAT
129301 ACAAAAAAGT AGCCAGGCGT GGTAGTGGGC GCCTATAGTC CCAGCTACTC
129351 GGGAGGCTGA GGCAGGAAAA TGTTGTGAAC CTGGGAGGCA GATCTTGCAG
129401 TGAGTCGAGA TCGCGCCACT GCACTCCAGC CTGGGTGACA GAGCAAGACT
129451 CCATCTCAAA AAAATAAAAA AAATAAAAAA AAGAAATATT ATGCTCAAAA
129501 TATATAGCAA TAAGTTGGAA ACTTTTACTT GAATAATTTT TACAAAACTG
129551 ACCAAAGAAC AAAAACGTGA AGAGGCCAAG TTCCAAGACT CATGTTATGT
129601 AAATTGTTCT AAAACAACAT TAGCACTTAA CAAAGTTGAA AAGTTAACAA
129651 AGCCAAGTAC TGTACTAGGC TTCCAACACT AACTAAGTAT AAAATTCCAC
129701 AGAGCTGGTT TTCTTATCTT TAAAGAAATT TGTTGGCAAG TGGTACTGGT
129751 GTTAAAAAAA AAAAAAAAAA AGGAAATATG TACTGACCAA AATAGAAAAA
129801 AAATATGAAG CACATTAAAA GAAAAAAATA TATATTCCTG AAAACCTTGT
129851 ATAATTACAG TGGCATGGTT GGGAATGTTT GGTCTATAGT TTTAACAATT
129901 AAATCCATTG AATCTGGCCC CGTACCATCC TAAAGTTTTA TTCTAGATTC
129951 TCTGGAGTTT GTGATTATAG ATATGTTTCT AAGATTTAAG TAACTTTCCA
130001 TGTTTATCTC CTTTTATGTT TGTACATAGA ATAAAAATGT TTCTATTGTT
130051 AAGAATATTA GAGTTGGACG CAGTGGCTCA CGCCTATAAT CCCAGCACTT
130101 TGGGAGGCCA AAGCAAGTAG TTTGTTTGAG CCCGGGAGTT CAAGAATGGC
130151 CTGGGCAACA TAGTAAGACC CCATCTCTAC AAAAAATAAA AAATTAGCCG
130201 GGCATAGTGG CATGTCCCAG CTACTTGGGA GACTAAAGTG GGAGGATCAC
130251 TTTGAGCCCA GGAGGTTGAG GCTGCAGTGA GCTATGATCG CACCACTGCA
130301 TTCCAGCCTG GGCAACAAAG TGAGACCTGT TTCAAAAATA AAAAATTGGG
130351 GTTTATCTAC TTAGATTTTC AATAAAAATT ACTACTTAAA TCTTTACCTG
130401 CTTGTTAATT TCAAACCCTT TTCTACATTT TGATTTATCT TTAAATCTCT
130451 TTTTGTCTCA ATAAATGGGA AGTATCAGGA AGTCTTTTTA CTTGCTCAAG
130501 GTCATAGAGA GCTTAGAACC TGGTAGTGTC CCTCTGAGCC CCAGTTCTTT
130551 CCAACCTGCC AGGCTGTAGG CCCAACAATT ACTCACCACT AAGAAATTAT
130601 GCTTGTGCTG TCATGGCAGT TGCATTGGAG AAAAGGATAT TTAACTGGCA
130651 AACAAAAGTC AGGAGAATGG GGAGATTTTG TTCTTTTGAA ATGCTAGTGT
130701 GAAGTGCTAG GCTTATTTTT CAAATGCCCA ACTCGTATTC TTTTCTTTTC
130751 TTTTTTTTTT GAGAGGGAGT CTCACACTGT CGCCCAGGCT GGAGTGCAGT
130801 GGGGCGATCT CGGCTCACTG CAAGCTCCGC CTGCTGGGTT GACGCCATTC
130851 TCCTGCCTCA GCCTCCGAGT AACTGGGACT ACAGGCGCCC ACCACCACGC
130901 CCGGCTAACT TTTTTTTTTT TTTTTGTATT TTTAGTAGAG ACGGGGTTTC
130951 ACCGTGTTAG CCAGGATGAT CTTGATCTCC TGACCTCGTG ATCCGCCCTC
131001 CTCAGCCTCC CAAACTGCTG GGATTACAGG CGTGAGCCAC CGCGCCCAGC
131051 CGGCCAACTC GTATTCCTAA ACGAATCATA ATTTTACCAT AAGACCATAG
131101 TTTAGTGATT GAAGAAAAAA TGTACCGAAC TGTATGATAT GATGGTGTCA
131151 AAAAGAACTA ACCCAATATG AAACAGTTTT CAGGAGCATG TTTCCTATTT
131201 TGGTGTCAGT GGACCACTTG TGTTAAAGTT GTGAAAACCA ACACTCCTGA
131251 ATTCCACCCA GAACTCACAC TCTGCACCTT CAGAGGCCCT CAGATTGTGA
131301 GTGGCTGCCC CGAGGTGTAC TACCAACCTC CAGCTTCCGT AGGTCCGTAG
131351 GTGTGCTAGT AGGGCCTAGG AAAAACAGAA CAGATGAGGA CAGTGATGCA
131401 TACAGCTGCT TTATCTGGTC TCTCCTTCCT CCCTAGCCTG ACTGCTATTG
131451 GAGGGCACCC TCCAGGCACA GTTCTGTTCA CAGCCTGCTG CTGCCCCCGT
131501 GGGCCATGGT TCCAGGACGG CTCCATCTTC TGTGCTTTGG GCACATTAAC
131551 CTCTCCCAGC GTCACTATCT TCATCAGCAA AATGGAGATA ACATTAGTAC
131601 CACCTCATAA AGTTGTTATG AGGATCACCA GTGAGATAAT CAATCTAAAG
131651 TGCTTACAAC AATGCTTGGC ACTTGGTAAA CACTAAATAA ATGATAGTTG
131701 CTATTATATG CATACTTTTA AAAAACCTGA TGCTTTTAAA AATTTTTTCT
131751 GCTGACTAGT GAATTGTTCA GTTTTTGTTG TTGTTGTTGT TGTTGTTGTT
131801 GTTTGAGACG GAGTCTCGCT CTGTCGCCCA GGCTGGAGTG CAGTGGCATG
131851 ATCTCGGCTC ACTGCAAGCT CCACCCCCCG GATTCACGCC ATTCTCCTGC
131901 CTTAGCCTCC CGAGTAGCTG GGATTACAGG CGCCGGCCAC CACGCCCGGC
131951 TAATTTTTTT GTATTTTTAG TAAAGACGGG GTTTCACCTT GTTAGCCAGG
132001 ATGGTCTCGA TCTCCCGACC TCATGATCTG CCCACCTCGG CCTCCCAAAG
132051 TGCTGGGGTT ACAGGCATGA GCCACCGTGC CCGGCCATGA ATGGTTCAGT
132101 TTTAACAGGT TCTGTGCCTT AAAAAAGTTA TTAAATTGAC TGTTTCCTCC
132151 TTTTTTGTAC CCATCATACT TTGAATATAT AACTACGGCA GCATATAAAC
132201 ACTTCACTTG CACTTATTTA TTTAAATGTC CATCTTTCCA TAACCCAAGG
132251 GTAGGAACGA AATCTTATTC ATTGTTGAAA CCTCTAGCAC ATAGCACAGT
132301 GCCTACCAAA TAGTAGGCAA AATCAGGTGT TCAATTCTAT TAACTACTCT
132351 AACACTGAAC TGAAAGTGTT CAATGGCTCA AAATAATATA ATAAGCCTTA
132401 ACTCTGGGGT GCTTAAATTT ATCTATAATC TCTGCCAGTG AAGTATACAT
132451 AGTTTTAAAG GTTAAAAAAA AATCAAGTGT TTAATGAATT TGAGCTGATT
132501 GAGCACTGAC CAGATACAGG ATCCTTAAAC TTCATAATAT CTAGTCCAAA
132551 GATGAATTTT TTTTTTGGTA CAGATTCTGA CTTCAAGGGA TTTGCAGGCT

FIGURE 3-39

132601 GGGAGGGAAA TTAAATATAT AAAAAGTCAT TTTCGGCCAG GCGCTGTGGC
132651 TCATGCCTGT AATCCCAGCA CTTTGGGAGG CCGGGCAGGT GGATCACTTG
132701 AGGTCAGGAG TTCAAGACCA GCCTGCCCAA CATGGGGAAG CCCCATCTCT
132751 ACTAAAAATA CAAAAATTAG CTGGACTTTG TGGTGCTTGC CTGTAGTCCC
132801 AGCTACTCAG GAGGCTTAGG CAGAAGAATT GCTTGAACCT GGGAAGCGGA
132851 GGTTGTAATG AGCTGAGATC ACACCACTGC ACTCCAGCCT GGGCAACAGA
132901 GACAGACTCC AGCTCAAAAA CAATACGTTT TTTTAATCTT GTCCCTTAAT
132951 GGAAATATTG AGAAAATGTC TAGGGGAACT CGAAGGAGAT GATTATAGGA
133001 GTTGATTATG TATGTAAAAT CAAAGTGAAT GAGGCAGTGG CAGGGGGGAA
133051 AGGGGGAACC AGTAATGACT TAGAAGTCCT AAGCATGTTG CATGGTAATT
133101 GTGACATTTG CTTCCTGCGA GCGGAGCTGA CCTTGTGGTG TCCGTCCTAG
133151 GTACTCTCCT TTTTCTTCAT TGGTCTGATG TCCATGATGA TTCCTCTGTG
133201 TAGCATCTTT GGGGCCCTCA TTGCTGTGTG CCTCATCATG GGTCTCTTCG
133251 ATGGATGCTT CATTTCCATT ATGGCTCCCA TAGCCTTTGA GTTAGTTGGT
133301 GCCCAGGATG TCTCCCAAGC AATTGGATTT CTGCTCGGAT TCATGTCTAT
133351 ACCCATGACT GTTGGCCCAC CCATTGCAGG TAAATATAAT GATTCTCCAG
133401 TAGTTATATT AATTCATAGT ATTTTCTACT TCAGGTCTTA ATTAAGTCTC
133451 ATTTATATGT AAAACATATT ACAGGTTATT GATTACTGGT CTTTTGCTTT
133501 TATGTGTGCC TTACTGGTAA CATTTTTAAT AAGACTAGCT ATTAAACAGT
133551 ATGTTGAATT TGCTGAAGAG TCCCTCTTAT CCTTACTGGC AGCTAAAACA
133601 CTATATAAAA CACTTTGGGA GGCTGAGGCG GGAAGATCAC CTGAGCTCAG
133651 GAGCTTGAGA CCAGCCTGGA CAACATAGTG AGACACTATC TCTACAAAAA
133701 ATAAAAAATA AAAAAATTAG CCAGGCGTGG TGGCATGCAT CTGCAGTCCC
133751 AGATGCACAG GAGGCTGAGG TGGGGGGATT GCTTGAGCCC AGAGGTCAAG
133801 GCTGCAGTGA GCTGTGATCA TACCACTGTA CCCCAGCCTA GGCGAAACCC
133851 TATCTTTAAA AAAAAAAAAA GTTGAAGATT CTGAAAGATT TCTAAATATG
133901 TCGGCTTTTG GTAAGAGATC ATACATCCCC CTGAGGAGTA GTTGATGAAA
133951 AATAACATGA TAGAAAAAAC CCAAATGCCT GTAAATAGCT AACTGGGTAA
134001 GCAAAATGTG GAATATCTAC AGTGCAATAT TATTCAGCCA TAAAAAGGAG
134051 TAAGGTACAG AACTGATGCT ACCAGGTGAA TGGACCTCAA GAACATTATG
134101 CTAAGTGAAA GAAGCCAAAC ATAAAAGGTG ACGTGGCTGG GCATGGTGGC
134151 TCACACCTGT AATCCCAGCA CTTTGAGAGG CCAAGGCAGG AGAATCATTT
134201 GAGCTCAGGA GTTCCAGACT AGCTTGGGTA ACATAGCAAG ACCTTGTCTC
134251 TACTAAAAAT AAAAAAAATT AGCCTGGCAT GGTGACACAT ACTAATATTA
134301 GTGTAGTCCC ACCTACTCCA GAGGCGGAGG CGAGAGGATC TATTGCGCCT
134351 GGGAGATCAA GGCTGCAGAG CCGTGATAGC ACCACTGCAC TCCAGCCTGG
134401 GTGACAGAGT GAGACCCTGT CTCAAAATAA AAGGTCATAT ATCGTATGAT
134451 TTATTTCATG TGAAATATCC AAAATAGGTA AATTCATACA GATAGAAAAC
134501 TGGTGACTGC CAGGGGCTGG AGAGAAAGGG ATTGGGAAGT AACTACTTAA
134551 TTGGAATGAA GTGTCCTTTT GGCATGATGA AAATATTTTG GAACTAAATA
134601 GAAGCAGTGG GTGCAAAACA TTGTGTATGT GCTAAATGCC ACTGAATTGT
134651 TCACTTTCAA ATGGTTAATT CTATTTCATG TGAATTCCAC CTCTAGAAAT
134701 TCTCTGAATG ATATTATAAA GATGTCCTAT GTTCAGTATT ACTTGGGTGC
134751 CTGTGTCAAT TAGTATTAGA AGATTCAGTA TAAATATGAT AAAATTTAAA
134801 TGTGTGTACT GTCATCTGTG TTTGGCACAT GAGGTGCTCA GTGAAGGTTT
134851 AGCTCAATTC ATTCTCATAC CCTGAACTGT ATTCCTAGTC ATGAATCAAG
134901 CAAGCAGAGT TACTACTGAG CAGACCAAAT CCAGTTTAAG TTCAAAAGGT
134951 CATGGTCTCC TTATAGGAGA ACCCTCCATT CCTTCTTCTC TCCCGCAGGT
135001 AATATAGTAT TCACTCTTCT TTCTATAGCC TACAAGCTTA GTTTTATAGT
135051 TTCTAAGTCC ACAGCAGATT TTATATAAGC AGCACTCTAG TAGTAAATAG
135101 TTTTGCTACC AGAAGTGACA GTTAACAGAG CAAGAATAAG GGAGGTAACA
135151 GGAAACTTCT AAAGCAGGCA GAACACCACC ACACAGCAAC TGGCAGCAGA
135201 AAGCTGCTGG CTCACCAGGA GGCTCTGAGG GCATCATAGT ATTACAGCAG
135251 AACATAATCA TGAGTGTTCA TAAGGAGGAA GAGTTCAGGA ATGGTCTGAT
135301 TCCAAAGATA GAAAACTTGT GCATTTACAG ATATTGTCTA TTTTATATAT
135351 GCAGAATGGA TGTACACATT TTTATGTGTT GGTACTTAAT TGGACAAATG
135401 GTCAGCTATT TATTTAACAA TGTCAGATAG ATATTTAGTA CAGTATATTC
135451 CCTGGACATT TTAGATAACT TAGGTTTTAT AGTATAAGTT ATAAGAGTTT
135501 AAAAACACTA AGATAACTTT TAAAACCATG AGTCCTGGAA TTTGTTAGAG
135551 AAATTAGAAA TGTTGAGTAC CGTAAAAGTT TTCAGAAGCA GAACCCGAAT
135601 ATAGAATGCC ATTAATAATT CTTATATACT TACATTTTCT TCCAGGTACT
135651 TATAAACCTT AGTCTATTTC CAAAATGTAT TTAACAACCT AACTTTCATC
135701 CAAATTATTA AATTTAAAAT TTTTCTTAAT TTTGAATTAT GATGTTAAAT
135751 AGTTTTCATG TTATTTCATA AAAAACAAAT ACACCAAAGA GTAGTTTGAA
135801 AGCTGGGCAT GGTGGCTCAT GCCTGTAATC TCAGCATTTT GGGAGGCTGA
135851 GGTGGGAGGA TTGCTTTAGC CTAGGAGCTT GAGACCAGCC TGGGTAATGT
135901 GACAAAACCC CATCTCTACA AAAAATATAA AAATTAGCCA GGCATGGTAG
135951 CATGCACCTG TAGTCCCAGC TACTCAAGAG GCTGAGCTGG GAGGATCAAT

FIGURE 3-40

136001 TGAGCTTGGG AAGTCAAGGC CTGCATGAGC CATGATTATG CTACCGCACT
136051 CCAGCCTGGG TGACACAGTG AGACCCTGTC TCAAAAAAAA AAAAACGTCT
136101 ACCTTTCTAT CTTTTCAAGT AATGTCTTGG TATGAAAATC CAGAGGACTG
136151 TACTTTATGA CAACGTTAAA GATATGAGAT CTTTTTCTTC CTATCAAAAA
136201 AGGTTTATAT TTCAGATCTG TTTCCTAAAA AAAAAAAAAA GTCTGATGTC
136251 TGAGATGTCT GAATCAGTCC TGTGGCTGAT CTAGACCCCC TACAGAGCCA
136301 CACTTGTTCT CCCTTGGTGA CAGCTTTTTC CCTTCTCAGG GTTACTTCGT
136351 GACAAACTGG GCTCCTATGA TGTGGCATTC TACCTCGCTG GAGTCCCTCC
136401 CCTTATTGGA GGTGCTGTGC TTTGTTTTAT CCCGTGGATC CATAGTAAGA
136451 AGCAAAGAGA GATCAGTAAA ACCACTGGAA AAGAAAAGAT GGAGAAAATG
136501 TTGGAAAACC AGAACTCTCT GCTGTCAAGT TCATCTGGAA TGTTCAAGAA
136551 AGAATCTGAC TCTATTATTT AATATCTTAC ATACCTCCAC CAGACTGGAC
136601 TTGCTTTTTG AATTTTAAGC AAGTTTCCTT TCCTTTTATA CAAATTGCAA
136651 ATTTCATATT TTTTTAATCA CATCCTAGGA ATAGCACAAT AATTGGGAAA
136701 TAGAACCCTT ATCACTAGAA GAACCATTTT CTGCCACTAA ATATCTCTGA
136751 TGTTTCCATG AGTCTGAGGG CAGAGACTCT GGTATATGAA AACATGTCTG
136801 AAAGTCACAT ATTGTGAAAA TTTGAAGCTA TCTCAGTAAA AAGCAGCTTT
136851 GGAAACTGTG AATGATCTTT AGCTTGTACA AATGTTTAAA AATACCTCAG
136901 GCTATACTGA AAGGGTTGCA GTTTGGTTAG GAGTGGAAAT ATTTTGTTTG
136951 TTAATGATGT CTTCAGTTCT GGTACCTCTG TTTTACTTTC TTATGCTCTT
137001 TGGAAACTTT TTGCAAAATT TAAGCCTGGG TTCTAGATAA TACCAGATCT
137051 ACCTAAACCT CAAGTCTATG TTAAAGTTGC TTTCCTGCTG TTAAATAAGC
137101 TATGATATTA AGATATTCTG ACTTGCTCCA GTGTCAAGGG ACCTTCTGGG
137151 AGCAGGTGCT AACATAGTGT TCAGAATCAA TATGTGAGAT GAAAAGGATC
137201 CCCTCCAGGA GGATCCTGAG CTGTTCAGAA ATCATTTAAG TTTACAGCGT
137251 TGTTCCCTTT GCGTTTGCAG TGCGTTTTAC TCAAGTAGCC AGAAACACCC
137301 CACGTTTCTG AATTTGTTTA AACTGTAACA ATAAAGTAAA ATAGAATGCA
137351 TGAAAGATAT TCTGGCGATT GTAACTTAGA ATTTTTCTGA CTTCTGGATT
137401 TGTTGGCACT AGAACCTGAT ATTTAAACAA AGTCTTACTG AGCAGCTATC
137451 AAGTGGCAGT TACAGGCACA AATTGGTGGA GGCTGGAGGA TGGGGAGGGG
137501 AGCAAAACCC TTTATATTTG TGAAGAAAAT ATCTGTAGCT GATAGAAATA
137551 ATTGCTTAAA TTGGTTTATG AAATTAATGA GTCTGAAAAG GTTAAAAGCA
137601 CTTATAAAAA GAACCAAGTC CTACATTTCC AGAACTTTCT GGCAAAAATT
137651 TGCACTCATA TTATTTATCC TATGAACATT CCCATTGTTT TTTTTTGCTA
137701 TTTATATACA GATTATCATA AGAAAGCTCT CAGTTTGAGG ACCCAAAATA
137751 AAACCAAAGT CATGCCATGA CCCATACTCA TTTACAAAAA CAAGAACACT
137801 TTCCTCTATC CCTAAAATTA TGCTTTAGTA CTTGAGGCCT TTAAAAGTTA
137851 GTGCTTTTGA TTGTGAAGAC ATTCAGCAAC TTACTTTGTC ATACATGCAG
137901 TTGCACCTTA CCACTTCTAA TAGTGTCATA TTTCATATTC AGGGGACTTA
137951 GATAATTTGC CTGTGGATGG TTCTTTTGCA GGAAAAAAAA TCTACATTTT
138001 GACCATACTA CCCTTTCATG TTCTTATTAT AAGCTTTTAG AAAATGATTT
138051 CATTCAGTCA TGTCCAGTTA TATAAAACGT TACTTTCTCA TTTTTGAGAA
138101 GTTCAACAAA ACATACTACT AAGACCAATC ATCAAACCCA CTATTATAAA
138151 TGTTAATTTT GGCTGGGTAA GGTGGCTTGC GCCTATAATC CCAGCACTTT
138201 GGGAGGCTGA GGAGGGAAGA TTGCTTGAGC CCAGGAGTTT GAGACCAGCC
138251 TGGGCAACAT AGCAAGATCC TGTCTCTACT AAAAATAAAA AAAAAAATTA
138301 GGCCAAGCAT AGTGGCTCAT GCCTGTAATC CTAGCACTTT GGTAGTCCAA
138351 GGCAGGGGGA TCACTTGAGC CCAGAAGTTC AAGACCAGCT TGGGTAACAT
138401 AATGAGACCC TGTGTCTACA AAAAATTTAA AAATTAGCCA GGCATGATGG
138451 TGCCCACCTG TAGCCCCAGC TACTCAGAGG CTGAGGTAGT GGAAGGATTG
138501 CTTGAGCCTA AGAGATGGAG GCTGCAGTGA GCTATGCCAC TGTACTCTAG
138551 CCTGTTCAAC TGAGCAGAAC CCTGTCTGTA AAAGAAAATC AAAAACAAAA
138601 AATAAATGTT AAATTTTGTT TTAAGTTTTA GCACAGACTC CCCTCAAAAC
138651 ACCTTCTCCC CAATTTTACA GAAAGTAATT CAAAAATGAA AACTTTACTC
138701 TGTAAAGACC TCTACAGTGT TTTTCTTTTC AAAATTTGGC TGATTTTAGG
138751 AAAAAAGTGA TCATCTGAAA CTAAAAGAAA TTGCTTGGTT AGTTTCCATA
138801 TTAAAACAGC AGTGACAAGT ATATATAACT TAGATCTCAG CATATGTGTT
138851 TGTATATTAA ACTTCACATA TGTAGTTTTC AGTTTAATGG AATGAATCAA
138901 ACTGGATCTA TAACACTGAA AAAGTTCTAT TGTAATAGAC TCATACGGAG
138951 AATACTCTGC TATAATAATA TAAAATTAAG AAGAAAAAGT ATAAACGTAA
139001 GATGCTAAAT TCCATAAATG CATATTTAGT ACTATGTTTT TTGTGGGAAA
139051 AGTTCTAAAA GTTTAAAATG CACAAAGAAA ATGAAAAATA CTAATATAAA
139101 AATTTGTGCT TTAATCTAGT CAAACTAAAT CCTTTCTAAT TTCTGAATGA
139151 AGTGTTACTG CTGCAATAAA GTGACCTGAT AAGCCTAAAT TTTTTGTGTT
139201 CAATCCAGAC ACTTTTCTGA GAGTCTGAAA AGAATACAGA GTCAGAACTC
139251 TGTTTTTATC TCCTCATCCT GTTTTTGATA AGACTCAGAA AATTCTCAAA
139301 TTCGAAAGGT TCTGGCATTT GAGGCCAAAA AAAGCATGAA AGGGAGTAAC
139351 ATTCCTTTTT ATAGATACTC TAGATTGGAT ACTATTGTAA CAGATGGCCA

FIGURE 3-41

139401 AGAAACTTCC AGAAACATTT TGGTTAAATT TTATTGCAAT GGATATTGCT
139451 GGGATCCATC CATTTAAGCA GTAATATACC ACCCAGATTA TTGATACTTT
139501 ATGCAAGATG TGTTCATCTC TTTGATCATA TTTACAATGC TTACTCCATA
139551 GCCCTGCTAC AAGACTTAAA ATT
(SEQ ID NO: 3)

Feature:
Start: 2104
Exon: 2104-2446
Exon: 87054-87198
Exon: 91571-92024
Exon: 120932-121075
Exon: 133151-133379
Exon: 136340-136569
Stop: 136570

Sim4 results:

Exon: 2104-2446, (Transcript Position: 1-346)
Exon: 87054-87198, (Transcript Position: 347-491)
Exon: 91571-92024, (Transcript Position: 492-945)
Exon: 120932-121075, (Transcript Position: 946-1089)
Exon: 133151-133379, (Transcript Position: 1090-1318)
Exon: 136340-136572, (Transcript Position: 1319-1551)

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
352	A	G	Intron			
381	C	T	Intron			
3505	G	A	Intron			
10280	G	C T	Intron			
11107	G	A	Intron			
15750	T	C	Intron			
16004	T	A	Intron			
16871	A	G	Intron			
17163	T	C	Intron			
17966	A	G	Intron			
19392	C	G	Intron			
20113	T	C	Intron			
20434	G	A	Intron			
21243	T	G	Intron			
23009	C	T	Intron			
24699	-	T	Intron			
28058	A	T	Intron			
29600	T	C	Intron			
31455	A	G	Intron			
35653	T	C	Intron			
42700	A	G	Intron			
45516	G	A	Intron			
51789	C	T	Intron			
52042	C	T	Intron			
52139	T	C	Intron			
53089	A	C	Intron			
53117	C	A	Intron			
53434	-	T C	Intron			
55431	T	G	Intron			
55905	C	T	Intron			
60567	C	T	Intron			
60751	C	T	Intron			
60755	G	A	Intron			
63301	T	G	Intron			
64573	T	A	Intron			
76462	T	C	Intron			
77652	G	A	Intron			

FIGURE 3-42

77819	G	A C T	Intron			
79594	T	C	Intron			
84331	A	T	Intron			
86107	C	T	Intron			
86175	A	-	Intron			
87109	C	T	Exon, coding	133	V	V
89444	A	T	Intron			
90535	G	A T C	Intron			
91163	T	A	Intron			
93488	A	-	Intron			
96065	T	C	Intron			
96351	C	T G A	Intron			
96701	T	C A	Intron			
96879	T	-	Intron			
97648	G	T	Intron			
97814	A	G	Intron			
98430	C	T	Intron			
101268	A	G	Intron			
103881	A	G	Intron			
103926	C	T	Intron			
107845	C	T	Intron			
109010	-	T	Intron			
109623	G	A C T	Intron			
110188	A	T C G	Intron			
111006	C	T A	Intron			
111223	A	G	Intron			
111457	T	C	Intron			
112168	T	C	Intron			
112653	G	-	Intron			
114155	-	A T	Intron			
114181	-	T	Intron			
114183	A	T	Intron			
115964	A	C	Intron			
118100	-	A G	Intron			
119631	A	G	Intron			
120833	T	C	Intron			
121125	A	G	Intron			
121245	C	T	Intron			
121521	G	A	Intron			
124296	C	T	Intron			
124549	G	A	Intron			
124858	G	T	Intron			
125920	A	T	Intron			
126266	A	G	Intron			
128258	G	T	Intron			
130303	C	A	Intron			
130617	C	A	Intron			
130910	-	T	Intron			
131727	C	T	Intron			
132895	G	A	Intron			
133506	G	A	Intron			
135473	G	A	Intron			
136201	A	G	Intron			
137080	A	C	Intron			
138022	T	C	Intron			
138543	A	T	Intron			
138681	C	T G A	Intron			

Context:

DNA
Position

352 GATGGTTAGCCAGGGATTATGGGTTTTGGCAGGAAGACCACAGAGGTAAAGTACCATTTT
CATCACATCATATCGGGGATACATTATCATCTAGTTGAGGTACTGTGTGCCATTTTTTGC
ACCCTAAAGTTATTTCTTCCCCCCACTCCCCCTTTCCATCCTATACTCTTTGGAAGAAAG
TTACTACGCATACCCACACTTAAAGAGTAAACCATTGTACTTCACCTCCATGAGGGAGGG

FIGURE 3-43

AGTATGTTCATAAAGTATTTACATTTCCTGCAGGAGAGATTTGTCTATTCTCTCCTCATT
[A,G]
TTTATTTAATCATTTACTTACATCAGTACTGACTCGTGGATAATTCTTACATATGTGTTT
GTTTGTGTGCATGCAAATATATAATCGATGTGCTTTCTTTGCCCAATAATATGTTGTGGA
CAACTTTCAAAGTCAATAAATACAGATGACCTTCAGAACTTTTAGAGGTTTTAAAGTAAG
TATCTAATCAGTCTTCTACCAATGTACATTATACTTCCAAATTTTCCTTATTTCCAACAA
TACTGGGGTATCATCTTCATACATACATTTTTGTGCACTTATGTGCCTATTCCTTTGTTT

381 CAGGAAGACCACAGAGGTAAAGTACCATTTTCATCACATCATATCGGGGATACATTATCA
TCTAGTTGAGGTACTGTGTGCCATTTTTTGCACCCTAAAGTTATTTCTTCCCCCCACTCC
CCCTTTCCATCCTATACTCTTTGGAAGAAAGTTACTACGCATACCCACACTTAAAGAGTA
AACCATTGTACTTCACCTCCATGAGGGAGGGAGTATGTTCATAAAGTATTTACATTTCCT
GCAGGAGAGATTTGTCTATTCTCTCCTCATTATTTATTTAATCATTTACTTACATCAGTA
[C,T]
TGACTCGTGGATAATTCTTACATATGTGTTTGTTTGTGTGCATGCAAATATATAATCGAT
GTGCTTTCTTTGCCCAATAATATGTTGTGGACAACTTTCAAAGTCAATAAATACAGATGA
CCTTCAGAACTTTTAGAGGTTTTAAAGTAAGTATCTAATCAGTCTTCTACCAATGTACAT
TATACTTCCAAATTTTCCTTATTTCCAACAATACTGGGGTATCATCTTCATACATACATT
TTTGTGCACTTATGTGCCTATTCCTTTGTTTACTATTTTACCCTCATTTCTAAGGCAGAT

3505 GGCTGGTTAGGCAAAAAACAAATCTAAGACCTTCTGCATGACACTTTAACATAAATTCTT
TCACTTTATCCTGCAAGGTGAGCGCGGTCAACCCCATTTGGGTGAGAAAACTGTAGCTCA
GTGAAAGTGTCTTGGTGGGTAGTAGAATGGCAATAAAACACATATCAACTGACTTCAAGG
GCTAAGTGATTTCCATTACTAAATCAACCTCCCTCCCCATCATTTGGGGTAACTTTATAT
GATTAATAGTCTTTTTTTTTAACCTTGATTTTCTATTATTTTTAGAGTGAATATTTCTTA
[G,A]
GTCTTTAGTATGCATATGAGGAATGGGCAAGACTGTAATAAATTCTGAGACAAAGGTAAT
GCTGGGTTATGCTGAGAGTTTTAAAACCTGACATAAATACTATTAAACTATTTGTATCAT
TCTGCAACTTACTTTTCTTCCATTCCGCATCATGTTTGTGACTTATCCACATAATACCTC
AGTGTGAACTGATAACTCAAATTCTTTCCATTTTAACTTAGGTGGTTTGCATTGTTTGAC
TATATTATACTCTATGCATTCTCCCTCTGATGGGCATTTAGATTGCTTCCAAACTCATTC

10280 TGGGAGGGAGGTAATTACCTGGTGGGAGGTAATTGAATCATGGGGGCAGGTTTTCCTGTG
CTGTTCTTGTGATAGTTAATAAGTCTCATGAGATCTGATGGTTTATAAAGGGCAGTTCCC
CTGCACACTTTCTGTTGCCTGTTGCCATGTAAGACATGCCTTTGCTCCTCTTTCACCTTC
CACCATGATTGTGAGGCCTCCCTAGCCATGTGGAACTGTGAGTCCATTAAACCTCTTTTT
CTTTATAAATTACCCAGCCTCAGATATTTCTTCATAACAGTATGAAAATGGAGTAATACA
[G,C,T]
TCCATTACCATAAAGAAAAGGCTTTCATGTACATTATTTTTTAGAGTAGCCTTGTGGTAT
GTCATTACCTCCATGGATAGATAAGAAAGTTGCAACTTGCACAGTATTAGGATTGATATC
AGTATTTACTTTTATTAAGTTGAACTTAAGAGCAGCTTTTTGGCTGGAAAAAAGTTGTAC
TTATGTCAAAGTTGTCCTGAAAGTAGAATCCTACTCCTGTCCCCAGCCTGAAACTATTTA
CTACATATTTACTTGCATGTTCTTTAGAATATTCTCTCAATAGTGTCTCCTACTCAAGTC

11107 GGAAACTCCAGCACACCTGGGAACTGCGCAGACCCACCACATAAGACAGATAGCCTATCA
GTGGCTGGAGGAATGGAGGAAAGCAGTGCTTTCAAATGTACATGCCAAATGTGTATGATC
ATACCTCTTTGTTAAAGTGCCTTCTTTAACAGCAAAAGTAATTCCTCACCTTGCATATAG
GAACTAAAAAAAAGTCGATGAAGAAATGGCTTGCCTTATTTTCAAGTAAGAAGTCTTTTT
TCATTTCACTAATTTTTAATTATGGGCATAAGTATGAAATACAGATTAGAAATACTGAAT
[G,A]
TGGACCAAAGCAATGTTTCCTTTGTGGACCAAAGCAGTGAATCTTTTTTCTTTCTTTCTT
TCTTTTCTTTTTTTTTGTTAAGAGACAGGGTCTTGCTCTCTTGCTCAGACTGGAGTGCAGT
CAGTGATGTGATGGCTCACCATAACCTCAACCTCTTGGGCTCAAGGGATCCTCCTGCCCC
AACCTCCTGAACAGCTGGGATTACAGGCACATACCACCACACCTGGCTAATTTTTAAAAA
TTTTTTTGTGGGGGAGGGTCTCTCTATATTGCCCAAGCTGGTTTCAAATGCCTGAGCTCA

15750 TGGCTCACTGCAGCCTCGACCTCCTGGGCTCAAGTGATCCTCCCACCTCAACCTCCCAAG
TAGTTAGGACTACAGGGGCATGCCGCTACACGTAACTAATTTTTGTATTTTTTTGTAGAG
ACAGGGTTTTGCCTTGTTGCCCAGGCTGGTCCTGAAATCCTTGGCTCAAGCAATCTGTCC
ACCTCAGCCTCTGAAAGTGCTGGCATTACAGGTCTGAGCCACTGCGCCCAGCCTAGATTT
TTTTGAATTGTAAAAAAGTAACCTGCTCCCTACTGAAGTAAATAGAGTTAAAAAAAGTAA
[T,C]
CTGGTACAGACACCTGTATTTTCTGACACCCCTAGAAGAGTCCCAGGTACCCTATAATCA
AATACATTAACATTTCTGCAGCAAAATGTATGGATAAGTGAGTTAAATAGAGACCATGAG
TAGCTTCAGGTCAGTTCAGATCAAGTTTTGCTTCTAATTAAATGTTGATATTCTCTTACA
AAAACTTTGGGTTTGGGTTTTCAGATTTGCAAATAAATAATTATAAATTATTATTTTTTT
TGAGACAGAGTCTTGCTGTGTTGCTCAGGCTGGAGTGCCATGGCACGATCACGGCTCACT

FIGURE 3-44

16004
AAAGTAACCTGCTCCCTACTGAAGTAAATAGAGTTAAAAAAAGTAATCTGGTACAGACAC
CTGTATTTTCTGACACCCCTAGAAGAGTCCCAGGTACCCTATAATCAAATACATTAACAT
TTCTGCAGCAAAATGTATGGATAAGTGAGTTAAATAGAGACCATGAGTAGCTTCAGGTCA
GTTCAGATCAAGTTTTGCTTCTAATTAAATGTTGATATTCTCTTACAAAAACTTTGGGTT
TGGGTTTTCAGATTTGCAAATAAATAATTATAAATTATTATTTTTTTTGAGACAGAGTCT
[T,A]
GCTGTGTTGCTCAGGCTGGAGTGCCATGGCACGATCACGGCTCACTGCAACCTCAACCTC
AGGCTGAAGCCATCCTCCCACTTCAGCCTCCCAAGTAGCTGGGACTACAGGAGTGTGCCA
ACATGTCCAGCTCATTTTTGTATTCTTAGTAGAGACAGGGTTTCGCCGTGTTGGCCAGGC
TGGTCTCAAACTCCTGGTCTCAAGTGATCCGCCTGCCTTGGCCTTCCAAAGTGTTGAGAT
TATAGGTGTCAGCCACTGGGCCTGGCAGAATTATACATTTATATGTCAATATTTGCTTTT

16871
GGGTATGAACAACTTTATAACACCTGTTACACATTGCCAGATTATTCTTTAGAAAACTTG
AATCAGTTTATTGTGCCACCAGTGATGTGCTGGCTTCCTGAAAACCCTACCAATGTTTGG
TTTTATTTTTATTAGTATTTGCTAATTTGATAAGTACTAATGATATTTTTTAAAAGTAGT
TTAAAATCATATTTCAGTGCTTATAAGTCTGTGTTCCCAGTTTTTGAGCCCTTTAGAAGC
TGCAAATGACCTGGCAATTATATATAATATTTGAAAATACAAGAGGACATATGCCAGTGA
[A,G]
TATATTAGAGTAAAACTTCATTCCCATAGGTAATGAAGGAATGCTTGAGATTATCTTAGG
CCTTAGATTCTCACCTGACACATCTTGGCAGGTAGACCATGTCCTTGTTTCCTCTGCTGT
CTTAGCCCAGGTGTTGATCAAGGTCTGTCTTAGGGCGGGGGATAGGAATGGAAATAAACC
ATGTAGAGACTTGGGCATGAGGACTTTGTGATTCTTCCAGGTGACATCTCATCCTTCAGA
GGATCAAGTCTGCAAGAGTAGCCATATCTTAATCTCTTTCAGTGCTATCACCTTGCATCA

17163
GCCAGTGAATATATTAGAGTAAAACTTCATTCCCATAGGTAATGAAGGAATGCTTGAGAT
TATCTTAGGCCTTAGATTCTCACCTGACACATCTTGGCAGGTAGACCATGTCCTTGTTTC
CTCTGCTGTCTTAGCCCAGGTGTTGATCAAGGTCTGTCTTAGGGCGGGGGATAGGAATGG
AAATAAACCATGTAGAGACTTGGGCATGAGGACTTTGTGATTCTTCCAGGTGACATCTCA
TCCTTCAGAGGATCAAGTCTGCAAGAGTAGCCATATCTTAATCTCTTTCAGTGCTATCAC
[T,C]
TTGCATCAACCTCTGGACTCGAGCTAATTCCGTTGAAAATATTTTATTAATTAATTTTGG
GGTATGTTAAAAATTTTGTTTGCATGTATTTATTTAGTTATTTTTATGAGACAGGGTCTC
GCTCTGTCACCCATGCTGGAGTACATACGGTTGCACGCTCATGGCTCACTGCAGCCTTGA
CTTCCCAGGCTCAAGTGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTACAAGTGC
ATGTCACCACATTTGGCTAATTTTCATATTTTTTGTAGAGACGGGGTTTCGCCACATTGC

17966
TCAAGCGATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGATTACAGACACACGCCACTAT
ACCTGGCTAATTTTTGTATTTTTAGTAGAAATGGGGTTTCACCATGTTGGTCAGGTGGGT
CTCAAACTATTGACCTCAGGTGATCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTATA
GATGTGAGCCACCATGTCCAGCCACCCATTTAATTTTTTGAGCACAAAATATGTACTGAG
AGCCACGCAAGAAACAAATTCGACTTATTCCATGCTCTTGAGAGGTCATGAGGGGAAACA
[A,G]
AATGATACATAAGTAACTCTGAGAGAATATGCTGCACATGCTAAATCCTGTGCAAGTAAG
ATATAGGATTTTAGAGGAAGGGAGAATGACTTCTGATTGAGCTGACTAGAGAAGGCTTCA
GTTTTTGAGTTAGGTGTTACGAGATTGGGAGACTTTTCTCAGCATATCTAACAGAAGAGG
GTATCCGAGGTGAGAGTGTAAGGCCTGGGCAAGGGTTGGGAGGCAGTTCTAATACTGAAT
GTTCTGACTGTGGTTTACTATGTATTTCAGGTTATTTTGTTTAATCTATCCAGTAATCCT

19392
GAACATGACATAATGGAAACCCAGAAATCCTGTCTTGAGCGGACTCAGGGGCTGCTTCTA
GGTAATTTAGTTCATTTCTACTGAAATCATTATATTTAAAAGTATGGCCGCACTGAGATG
GCCACTGTAGCTGCTGCCACCTCTTAGCTTTGGTCTTAAAAAAAAAAAAAAAGTAATAGAA
CTTCCTTAAAATGTCTTTCTAGCCTTTGGATTTCTCTAATTCAGATTTGCGTCTTCCCAA
GGGTCAAAATTATATTTTTACTATCCCTGTCTTAGGTATTTCCAAAACTTCGTCTTAAGA
[C,G]
TTAGTCATTTTTTTCCTTCAATTTGACATGACTGCTAAAGACTTTTGGCATGTTCCTCCT
CCTTTCATTTGTGATGTAATTAAGTTGGTCTGTAAGTCTTATTTTTAAGATGTTCTAGAC
CAAGAGACTGTGAGAATAGCTTACAGTCATTTCAACTAATTTATGTATTTTAAATTTAAA
GTATTGACAGTGGTGAAAACCTGTTCAACAAGCAGATGATGTTATCTTATATATTCACAG
AGTTTAGTAACTGAGCCAACTACTTCATTCACAGTTCAAAATGAAAACAGCTAATTCTTT

20113
TTGAGAAATATGGACTGTTTTTGCACATTCATAATGGACATTTGAGGTTTGTTAGGGGAG
GAGGTGTCATCTTTATGGCACTTTCTGGCTGGGAAGGGAGTCAGTCCTAATTGAGATAAT
AACTAGCCACCTGGCCACACACAAGTGTGTTTTGCCTTAGTTACCTGTCACACACTGAGC
AGTGAGACTCAAGAGAGTGTCAAAGTACTTTTCAATGCATAAAGCACTACAGATCTGTCC
ACACTGTTGTGAGTGAGCAGGTTGGCACGGTGCCTGTGTGCGGGGCGTGTGTGTTGACTCA
[T,C]
GTGCTGTCCTGTGATCTCTCAGGACTCAGGTCCTGAATTGCTCTGTTGTGACTGAAGCCC
AGCTGAAGGTGCTGGAAGTGCAGTGACCCTGGAGGAAGAACCAGTAACAGCAGAGGGTGG

FIGURE 3-45

```
ATGAAAGGGAATTGATAGTTGTGTAAGAATAGATATCTGCCGTTTTTTGTAAGCCAAGAC
ACCTTTACCCTTCCAGTAATTGTTTCATCTTTTAATATCATTTGGCTTCATTTACAGAAT
CTGTATTAAAGCAAAAGTCAGCATGTAAGGTGGTATTTTGACCACATTTGTCTGCTGTTG

20434 AGGACTCAGGTCCTGAATTGCTCTGTTGTGACTGAAGCCCAGCTGAAGGTGCTGGAAGTG
CAGTGACCCTGGAGGAAGAACCAGTAACAGCAGAGGGTGGATGAAAGGGAATTGATAGTT
GTGTAAGAATAGATATCTGCCGTTTTTTGTAAGCCAAGACACCTTTACCCTTCCAGTAAT
TGTTTCATCTTTTAATATCATTTGGCTTCATTTACAGAATCTGTATTAAAGCAAAAGTCA
GCATGTAAGGTGGTATTTTGACCACATTTGTCTGCTGTTGTGCCTTCTGGGTGAACTGAA
[G,A]
TACTGGCTTACTGACTAGTAAATATGTTTTCTGACAATTATAGGGAAGGGAAGAAAAGGA
AAGTCCAATTAAAGCATTTTTCTCCTCAGAGTTTGAAAAATAGAATTCATGCAATCTTTT
AAATTCCATGCCAACACATCAGACAAGAAGAGACTTGATAGTAGTAAAGGTTGGGAATCA
AAAGAACAATGTAAGTTTTTGATATTGACTTCAAAACATGGTGTTCTATAATTTAGTGTT
CATTTGTTACGTGTATGGTATTATAATTAATTTTGTATATGTGGTAGTTATTTTTTGTAC

21243 TAAATTAATCAGCGGTCACTAATCTTTGGATAATCACTCTATTGAGCTGGAACTATCCTT
AGTATTTTGGAAAGCAAGTCAGTGAGTTAGAACTGTCAAAACTGATCAGCTTTTCTAAGC
TTAATGATAAGTGAATAGAAACTAGTTGCCTTCAACCCTTTCCTCCCTGCATTGCAGCAT
GATCATTCTGTAACTCTGGAAATGGTTTATGGAACAACAGTGAAAATACATTGATACACT
GTCTTGTGGTAGATTTTCAGATAGGCTTTAGACAAAGTTCAGAGCCTTTCCTCTAGCTGG
[T,G]
GATTAACAAAGCTGCCTTCATAGTTAAATGTTTGCACCCTGTGTATGCATTTTCAGTTAC
TAGAATTAGGTAAGTTAGTGTTATAAAATTGGTTTGAGTGTGGATTGTTTAGGAAGTGAG
TCTTTTGGTGGCAGCAATTCTGTTATGCATTAAATAGATACATATTTTGAACTAGTCGAC
ATTGTTTCAGTCTGTATTTATTAGATGCTGGGGTGGGTATGGGAATAAAGAAACGTATGA
GGGGTCTTGGAAAAGTTCATGGAAAAAATGTACACTATGAAAAAAAACTGTGCATGGATT

23009 TTCTTTCCTAATCTTAAAAAATATTTAAAGGGCACCTATTTTTCTTCAGTTAATAATGTA
AAAAGGACTGCATTGACATGATTAAATTCCTGGGACCCTCAATTCTTTAGAGATGGACTA
ATGGCTGGTATCAACTCACAAAAGTATCTTGAACTTGATGGAGCTTATGTTGAGAAATGA
AGTGTATATTTTCATTATCTTTTAATTTCATTCTTTAGTGAATTTTTTGAGGTCCCCTTG
TATACATTTTAATCCTAAGGGAATAAAGAAAGGAGGAAGTCCTAGCCCTGTGCTGTCTGC
[C,T]
TAGGTACAGTGTCTGAAACACAGACCAGTATTCACCCTTTGAAATTTGAGGTTTCCATTC
AGGAGGTTCTCAAAGAGAATAAATGAGATTGCTATGCAGGTGGAATCAAAGAGCACACGG
CTTATTTATCATAATCAAAATAATGCCATTTTCATAACAAACTTCACCTGCTTATGTACA
TTGTAAATTGTTGCCTTGATAAGCTTCCCGGAGATAAAGTAATTCAGCTAAGTATTATTT
CCAATCATAATTTTGTGTCATTATGAGCAACACAATACTATATATGGGATTGATTCACTG

24699 TTTAGAGGATTGGATGAATAGTGGTGCTGCCAATAAAGAAATTTAAATATGGCTGATATT
TCCTATATTTAAGAAAGACCAAAGAGGGTCCATTGAAATGAGTCAGTGGGAAATCTCTGA
TGACTTCAGCCAGCACGCTTTCATCGGCCTGGATATATGGGAAGTGAAGTCTGATTATAG
TCTGTGGAGCAGTGAATGGGAGGAAGAGATAGGGGTACAGGCTAAGAAGGGAGGAAGTCA
AGTCAAAGGGAGAAGTAGGGTGGTAGCTAGAGGAAGATTAGAGTCAAGCGAGGGTAACAA
[-,T]
TTTTTTTTTTTTTGAAGATAGGAGTAGCTTGAGAACTAACTTAAAGAAGGAGCCTGTAGA
GAGGGAGGAGGTGAAGTTACTAAAGGTCTAATTGATGGGGTAAGGTTCATGAGCAGATCA
GATCTTTACAAGGAAGGTTCTTGCTGGGGGGCAAGATTCAAAACCCCTATTCAGAACCAG
GAGAGAAGAAAGTAAGAATGGGAGCAAATGTAGGTAGGTTTGGTGAGGATCAGGAAATGG
AGGGGAAGAGGTCATTAAATGTGGTCCTGGGGTTGAGCAGCAGATTGGAAGAGAATGGCA

28058 GCAAGGTGTTTTTCGATCACATGGAAAGGGGAAGAAGAAGCTATCAGGAGCTCTGGGGTT
TTTTTTGTTGTTGTTTGTTTGTTTTGCCACTTTTAACTCTCAAGCTAAAACTGGGGTTTC
ATTTGAGGAACCAGTAATAGAAAATTTCTTATGTACATTCAGCAAAATCTAGTACTGAGT
GGTTACTTTGGCTTTTCATTGTGGGGATTGTGTGTGTGTGAGTACATGCACGCACTTGTG
TGTTTAAGCGTGTAAGGCAGACAGACAGTGGGTACAGGTCTTTGAAATGGACTTCTTGGC
[A,T]
AAAGTAATAGAGAAAAAGAGGAATACAAATAAGGGAGGAGGGACAGGGAAGAGCAGAGTC
ACAGGAAACAGTGAATGAGGCTGCAGTCTCAGTCGCCCTTTCTTTGTCCCTCCAGTGTTG
TTGCCTGTCTTATGATGATGCTGGTTTTCAGCCAACCTTGAGTGAGTAAAAGCCGGGTCT
GAGGTCTCAGTGCCTGCGTGGCTGATATGAGCAGCTTGCATTTCTGACTGGGCCCTGGAG
CAGCAACAGCACAGATTTCCAGGAACAGTTCCTCTTGTCATTTTTATTCCTGAGTCATCA

29600 ACTGCAACCTCCTTAAAGCTACATTATTTAAAAGTCACATACAAAGCAAGTTGCAGAAGC
CTGTATGTAGTGGATTCTATTTTTTTTAAATAGTATTTAATTGTATGTTCTTCTACACTT
TTTCCTATGTCCATCTTACCATAGCTGTGCCTTTTTTGGTGGAAGTGAGGACAGATTGCT
TTCCACATCTCCATTTTTGTGTCTGAATTAAAAGATGGACAAGTATCATGTATTATCTTA
```

FIGURE 3-46

GTAGTCATCAAACAAGGAAAAAGGTTTCTTTGTTTGCTTGTTTTTTTTAGATGAAGTCTC
[T,C]
GCCCAGGCTGGAGCGCAGTGGCACGGTCTTGGCTCACTGCAACCTCTGCCTCCTGTGTTC
AAGCAGTTCTCTGCCTCAGCCTCCTGAGTAGCTGGGATTATAGGCGCCTGCCATCACGCC
GGCTAATTTTTGTATTTTGAGTAGAGACAAGGTTTTGCCAAGTTGGCCAGGCTGGTCTTG
AACTCCTGACCTCAGGTGATCCACCTGCCTTGGCCTCCCAAATTGTTGGGATTACAGGCG
TGAGCCACTGTGCCCGGCCTGGAAAAAGTTTTTAATGGTAAAGATGTCATGGAATGAATA

31455 ACTTAATTCGTTGTGGGCAGCCAGATCTTTTAAAGGTAAATTTGAATTTCTCTTTAAGAA
AATGGCAGACAGAAGGATGGGGGATACTAGAAAACTAAAAGTAGTGCCCCCTTTGAAGAT
AAAACTAAAACATTTTAAGCCTGGAATTGCTTTAGCAGTACATGTATTGATTATTTAATT
TTGTCCTTTAGAAGAAAGTTGGCCCAACACAATTACATGGAAGTTGGGTTATTGAAGAGG
ATTGATAAAAGAAGTGGGAAGGTCAGGCCAGGTGTGGTGGGTCATGCCTGTAATCCCAGC
[A,G]
CTTTGGGAGGCCGAGGTGGGTGGATAACGAGGTCAGGAGATCGAGACCATCCTGGCTAAC
ACGTTGAAACCCCGTCTCTACTAGAAATACAAAAAAAATTAGCCCAGCATGGTGTTGGGT
GCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCG
GAGGTTGCAGTGAGCCCAGATCAAGCCACTGCACTCCATCCTGGGCGACAGAGCGAGACT
CCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTGGGAGGGTCAAAGCCAATGTGCACGT

35653 CTAATAGAAAAAAATATATATTTTTGTCAAATATTTCTGTTTTATTCATTCATTCAAAGT
ATATTTGGAATGTTTATTTCCAGGAAATTTTGGAATATACAATACAACCAGCTTCTTATA
ACTCCACTTTAAGTGAGCCATAGGTCAAATAATGACCAGCAAAATGTAATGACACGTGTG
CCTCTTACTCCCTGTTGGAGGAATTGAGGCACTCTGGTAACCCTGTAGGCCTGGATTAGT
CCAGTTCATTGGCAGCAGCATTATCCAGATTTTATTGTGGCCGGCAACGGTGGCTCACAC
[T,C]
GGTAATCCCGGCACTTTGGGGGGCTGAGTTGGGCCTGTTGCTTGAGCCCAGGAGTTCAAG
ACCAGCCTGGGCAACATAGGAAACCCTGTCTCTACAAAATATATAAAAATTAGCTGGGCG
TGGTGGCGTGTGCCTGTAGTACCAGCTACTTCGGAGGCTGAGGCAGGAGGATCACCTGAG
CCCAGAAAGTTGAGGCTGTGGTCAGCTCTGATTATGCCACTGCACCCCAGCTTGGGTGAT
ACAGTGAGACCCTGTCTTAAACAAACAAAAGAGATTGTATTGTGTTTTGAAAAACATAGT

42700 AGCTTTTCTCGTCTCTACTGAGGCCAAAAGGGGCAGTGATACCCTTGAATTTTCTTCTTA
AAACAGGGTTTCATTTCCTTGGAAGTTTGTTTCTTTGAATCTTTCTGTCAGTTTAACTGT
TATCATCAATTGGTTAGCATTCTAATAATAATTATAATTATAGTAAACATTTATTGAGTG
CTTACGAAGAGCCAGTTCCAAGCTTTTTTATCTCCATTATTCTGCTACTTTCCTTCTCAT
TTTACAGATGAGGAAAATGAGGCACAGAGTGGTTAATTAATCTGTTTGAGGTCCCGTAGC
[A,G]
GGTCAGTGATGCCAGGGTTCAAACCTACACTTAACTCTACACTAGAGACTGTTTTCTTAA
TTATTTCTTCACAATCATATGTTTAATGATTACTTATTGATTATTTAGTGGTCTGATAAG
AAGAGGGAGCGGTGCTCTTCTGTTGGAGAAGAAAGGCTGGCTGATCAAGACACACTGGTT
GGTTTGAAGAAAAAATATAGATGTTAATTCCATAACACCACACTCTAAACATTTCTACTG
GACGAGTTCCACCTGTGTGCCACTCGAAGTCGGATGCAGTAAGGAAGGCTTTTTATTGAG

45516 GTTACTAGCTATGTAATCTTGAGAAAACTACTCAATCTCTCTGTGCCTTAATTTTCTTAG
GTATAAAGCAGATACTAATTATGCCATCATAGGGTAGGTATGAGGATTAAATGAGTGAGT
ATTTGTAAAACACTTAAAACAGTGACTGACTTGGGCTACCCCTTTTGGGTCCCCTCCCTT
TGTATGGGAGCTCTGTTTTCACTCTATTAAATCTTGCAACTTCACACTCTTCCAGTCTGT
GTTTGTTATGGCTCAAGCTGAGCTTTCGCTCGCTGTCCACCACTGCTGTTTGCTGCCATC
[G,A]
CAGACCCGCCGCTGACTTCCACCCCTCTGGATCCGGCAGGGTGTCCACTGCACCTCTGGT
CCAGCGAGGTGGTGCCCATTGCCGCTCCCAATCGGGCTAGGGGCTTGCCATTGTTCCTGC
ACGGGCTAAGTGCCCTGGGTTCATGCTAATTGAGCTGAATAGAGCTGTAACACTCACTGT
ATGGCCCAAGGTTCCATTCCTTGGAATCTGTGAGGCCAAGAACCCCAGGTCAGAGAAGAA
GAGGCTTGCCGCCATCTTGGAAGCAGCCCGCCACCATCTTGGGAGCTCTAAGAACAAGGA

51789 AGAAGAACCGCTAGTATGGGGTAATCCCCTCCAAGAAACCAAGCCCCAGTACTCAGAAGA
AGAAATAGAATGGGAAACCTCATGAGGACGTAGTTTCCTCCTCAGGATGGCTAGCCACCA
AAGAAGGAAAAATACTTTTGCCTGCAGCTAACCAATGGAAATTACTTAAAACCCTTCACT
TAGGCATTGATAGCACCCATCAGATGGCCAAATCATTATTTACTGGACCAGGCCTTTTCA
AAACTATGAAGCAGATAGTCAGAGCCTGTGAAGTGTGCCAAAAAATAATCCCCTGCACTT
[C,T]
AGGCCATGCATTTCAATCCCTGAATCTTTAACCTCCTTGTTAAGTTTGTCTCTTACAGAA
TTGAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGATCCCCAGATGCAGTCTATGACTC
AAATCTACCGCGGACCCTTGGACCGGCCTGCTAGTCCATGCTTCGATGTTGATGATATCA
AAGGCACCCCTCCCGAGGAAATCTCAAGTGCATGACCCTTAGTTGCACCAGTTCAGCAGG
AAGCAGTTAGAGCGGCCGTTGGCCAACCTCCCCAATAGTACTTGGGTTTTCCTGTTGAGA

FIGURE 3-47

52042
GATAGTCAGAGCCTGTGAAGTGTGCCAAAAAATAATCCCCTGCACTTCAGGCCATGCATT
TCAATCCCTGAATCTTTAACCTCCTTGTTAAGTTTGTCTCTTACAGAATTGAAGCTGTAA
AGCTACAAATGGTTCTTCAAATGGATCCCCAGATGCAGTCTATGACTCAAATCTACCGCG
GACCCTTGGACCGGCCTGCTAGTCCATGCTTCGATGTTGATGATATCAAAGGCACCCCTC
CCGAGGAAATCTCAAGTGCATGACCCTTAGTTGCACCAGTTCAGCAGGAAGCAGTTAGAG
[C,T]
GGCCGTTGGCCAACCTCCCCAATAGTACTTGGGTTTTCCTGTTGAGAGGGGTTGCTGAGA
GACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTG
ACTGCATCCACCTTTAAACACGGGGCTTGCAACGTAGCTCACACCCGACCAATGAGGTAG
TAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAAGTAAAGAAATAGCCAATC
ATCTATCACCTGAGAGCACAGGGGGAGGGACAATGATCAGGATATAAACCCAGGCGTTCT

52139
CTCTTACAGAATTGAAGCTGTAAAGCTACAAATGGTTCTTCAAATGGATCCCCAGATGCA
GTCTATGACTCAAATCTACCGCGGACCCTTGGACCGGCCTGCTAGTCCATGCTTCGATGT
TGATGATATCAAAGGCACCCCTCCCGAGGAAATCTCAAGTGCATGACCCTTAGTTGCACC
AGTTCAGCAGGAAGCAGTTAGAGCGGCCGTTGGCCAACCTCCCCAATAGTACTTGGGTTT
TCCTGTTGAGAGGGGTTGCTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAA
[T,C]
GCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTTTAAACACGGGGCTTGCAACGTAG
CTCACACCCGACCAATGAGGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAAC
AGGAAGTAAAGAAATAGCCAATCATCTATCACCTGAGAGCACAGGGGGAGGGACAATGAT
CAGGATATAAACCCAGGCGTTCTAGCCGGCAACGGCTACCCTCTTTGGGTACCCTCCCTT
TGTATGGGAGCTCTGTTTTCACTCTATTAAATCTTGCAACTGCACAAAAACCAAACCAAA

53089
GATAAGTGAATCTTCAGAGAACTGGCCTTGAGCCAGCTCTACAACTAACCAGCTCTGTGG
CCCTTTGGAGAATTTCTTAATATTTGTAAACCTCAGCTTTCCTACCAGTGAAATGAAGTT
AGTCCTCCCTGTCCTGCAGGGTTGCTGCAAGGATTTAACAACATGTATATGTACAAACCA
CTTAGTCCTGTGCTTGGCCTATTTGGTGCTTTTTTTTTTTTCTTTTTTTTAAGACAGGGTC
TTGCTTGAATCTTGCTGAGGCTGGATTCAAACTCCGGGGCTCAAGTGATCCTCCTGCCTC
[A,C]
GCCTTCCAAGTAGCTGGGACTACAGGCCTGCACCACTGTGCCTGGTGGCAGTGCTCGTTG
AATGTTCTTTTTCCTTAGTTCCTTCCTAGCTCTTCTGACAGTTTTGGGGCTTATGTATAT
AAGAAGGACTTGGTTGCCTCAGGGAGAGAGGATGCAGTAGAGTTACATAGCTCACCTCAC
ATCCTCCAAAAGCTGAATTCATAAGTAAACAAAGTGAGCATTTCACCCATACTTTACACA
AAGTCTAGAATATTTATGGTGTCCATCAGGCTCACATACTGTGACCTTCTGAGATACTTT

53117
TGAGCCAGCTCTACAACTAACCAGCTCTGTGGCCCTTTGGAGAATTTCTTAATATTTGTA
AACCTCAGCTTTCCTACCAGTGAAATGAAGTTAGTCCTCCCTGTCCTGCAGGGTTGCTGC
AAGGATTTAACAACATGTATATGTACAAACCACTTAGTCCTGTGCTTGGCCTATTTGGTG
CTTTTTTTTTTTCTTTTTTTTAAGACAGGGTCTTGCTTGAATCTTGCTGAGGCTGGATTC
AAACTCCGGGGCTCAAGTGATCCTCCTGCCTCAGCCTTCCAAGTAGCTGGGACTACAGGC
[C,A]
TGCACCACTGTGCCTGGTGGCAGTGCTCGTTGAATGTTCTTTTTCCTTAGTTCCTTCCTA
GCTCTTCTGACAGTTTTGGGGCTTATGTATATAAGAAGGACTTGGTTGCCTCAGGGAGAG
AGGATGCAGTAGAGTTACATAGCTCACCTCACATCCTCCAAAAGCTGAATTCATAAGTAA
ACAAAGTGAGCATTTCACCCATACTTTACACAAAGTCTAGAATATTTATGGTGTCCATCA
GGCTCACATACTGTGACCTTCTGAGATACTTTTCCCTCTCCATTCCCTTTCCTTCTCCCT

53434
GTGGCAGTGCTCGTTGAATGTTCTTTTTCCTTAGTTCCTTCCTAGCTCTTCTGACAGTTT
TGGGGCTTATGTATATAAGAAGGACTTGGTTGCCTCAGGGAGAGAGGATGCAGTAGAGTT
ACATAGCTCACCTCACATCCTCCAAAAGCTGAATTCATAAGTAAACAAAGTGAGCATTTC
ACCCATACTTTACACAAAGTCTAGAATATTTATGGTGTCCATCAGGCTCACATACTGTGA
CCTTCTGAGATACTTTTCCCTCTCCATTCCCTTTCCTTCTCCCTGCTGGCTTTTTTTTTT
[-,T,C]
TTCTTCTTCTTTTTTTTTTTTTTTTTACTGTGAAAAACAACCTATATACAGAATAGTACAA
AAACATACCTGTATAGTTTGAAGAGTAATTATTAACAGTCTTATTAAGAAACAATGCTCC
ATCCATGTTACTGCAAAAGACATGACCTTATTCTTTTTCTTCTTAATTTTTTTCTTTTTC
TTTCTTTATTTTGGCCCTTTTTCAGATCTAGACCTGCAGAGATCTTGTTCTTTTTTTGAG
ACAGCATCTCGCTCTGTCACCAGGCTGGAGTTCAGTGGCGTGATCTCGGCTCACTGCACC

55431
TAATTATTTTATTTTATTTTATTTGAGAGAGGGTCTTTCTGTGTCACCAAGGCTGGAGTG
CAGTGATGCAATCATGGTTCACTGCAACCTCAACTTCCCGGGCTCCAGTGATCCTCCCGC
CTCAGCCTCCCAAGTGGTTGGGACTACAGACATGTGCCACCAAATCCAGCTAATTTTTAA
ATTGTTTTTAATAGAGGTAAGGGTCTCACTATGTTGCCTAGGCCAGTCTCGAATTCCAGG
GGCTCAAGGGATCCTTTTGCCTTGTCCTCCCAGAGTGCTCGGATTTAAGTTGGGAGCCAC
[T,G]
ATACCCACCCAACATAATTCAATTATTTAATATTTCACATGTTTTAGTATTCCTTTGATA
GGGATGTGATGTTTGGGTGAATAATAAAGTAAATCAAAGACATATATTTGAAAATTATGT

FIGURE 3-48

AGTTATTCTAAAAAATTAATTATTTACCTTTATTTTAGCAAAATCAGTGTGTTAGCATAA
TCAAGATATTTTGGTATTCTAGTAACAAGATCTAGTCACAGTAATGATGTAAAGATTAAA
AAATAAAATATAATAGGAACCAGTATAAACAAGTGGAATTTAAATTTAAAATGCAATACC

55905 AGCATAATCAAGATATTTTGGTATTCTAGTAACAAGATCTAGTCACAGTAATGATGTAAA
GATTAAAAAATAAAATATAATAGGAACCAGTATAAACAAGTGGAATTTAAATTTAAAATG
CAATACCAGCTGGGTGCGATGCCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGC
AGGCGGATCACCTGGGGTCAGGAGTTCAAGACCAGCCTGACCAACATGGAAAAACCCCAT
CTCTCCCAAAAATACAAAATAAGTTGGGTGTGGTGGTGCATGCCTGTAATCCCAGCTACT
[C,T]
GGGAGGCGGAGGCAGGAGAATCACTTGAACCAGGGAGGCAGAGGTTGTGATGAGCCGAGA
TCACACCATTGCACTCCAGTTTGGGCAACAAGGGCAAAACTCTGTCTCAAAAACAAAAAG
AAACAAAAAACACAGTACCATTTACATTAGCACCCCTCAAAATGAAATACTTAGGTATAA
ATCCAGCAAAATAGGTATAAGAGATATATAAGTAAAACTATAAAGCTCTGATGAAAGAAA
TAAAAGAACCAAATAAATGGACAGATATTCCATGTTCATGGATAGGAAGACTCAGTAATG

60567 TAAATATGACATTAAAAGCACAAACAACAACAAAATAGATTAATTGGACTTCATCAAAAT
TAAAACCTCTGTGCTTCAAAGGGCACACCAAGAAAGTGAAAAGAGAATCCACACAATGGG
AGATAATTTTTTGCAAATCATGTATTTTACAAGACTGGTGTCCAGAATATATAAAGAACA
CTTGCAACTCAGCAATAAAAAGACAAGTAACACAATTTAAAAATGTTGAAAGGATTTGAA
TAGACATTTCTTCAAAGAAGACATATAAATCACCAATGAGCATATGAAAATGTACTCAAC
[C,T]
TCATTGGTCATTAGAGAAATGCAAATAGAAGTCACACCCATTAGGATGGCTAAAATAAAA
AAAGATGAACAATAACAAATGTTGGCAAGTATGTGGAAAAATTAGAACCCTCATACACTG
TGGATGGGAATGTAAAATGGTGCAGACACTTTGGAAAGTTGGCTATTCCTCAGAGATTTA
CCACATGGCACAGCAATTCTACTTTTAGGTGTATACCCAAGACAATTAAAAAGATATATA
CAGGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGCCAAGGTGGGTGGATC

60751 CAACTCAGCAATAAAAAGACAAGTAACACAATTTAAAAATGTTGAAAGGATTTGAATAGA
CATTTCTTCAAAGAAGACATATAAATCACCAATGAGCATATGAAAATGTACTCAACCTCA
TTGGTCATTAGAGAAATGCAAATAGAAGTCACACCCATTAGGATGGCTAAAATAAAAAAA
GATGAACAATAACAAATGTTGGCAAGTATGTGGAAAAATTAGAACCCTCATACACTGTGG
ATGGGAATGTAAAATGGTGCAGACACTTTGGAAAGTTGGCTATTCCTCAGAGATTTACCA
[C,T]
ATGGCACAGCAATTCTACTTTTAGGTGTATACCCAAGACAATTAAAAAGATATATACAGG
CCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGCCAAGGTGGGTGGATCACGA
GGTCAGGAGATCGAGACCATCCTGGCTAATACAGTGAAACCCCATCTCTACTAAAAATAC
AAAAAATTAGCTGGGCGTGGTGGGGGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAG
GCAGGAGAATGTCGTGAACCCGGCAGGCGGGGCTTGCAGTGAGCCGAGATTGCGCCACTG

60755 TCAGCAATAAAAAGACAAGTAACACAATTTAAAAATGTTGAAAGGATTTGAATAGACATT
TCTTCAAAGAAGACATATAAATCACCAATGAGCATATGAAAATGTACTCAACCTCATTGG
TCATTAGAGAAATGCAAATAGAAGTCACACCCATTAGGATGGCTAAAATAAAAAAAGATG
AACAATAACAAATGTTGGCAAGTATGTGGAAAAATTAGAACCCTCATACACTGTGGATGG
GAATGTAAAATGGTGCAGACACTTTGGAAAGTTGGCTATTCCTCAGAGATTTACCACATG
[G,A]
CACAGCAATTCTACTTTTAGGTGTATACCCAAGACAATTAAAAAGATATATACAGGCCGG
GCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGCCAAGGTGGGTGGATCACGAGGTC
AGGAGATCGAGACCATCCTGGCTAATACAGTGAAACCCCATCTCTACTAAAAATACAAAA
AATTAGCTGGGCGTGGTGGGGGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAG
GAGAATGTCGTGAACCCGGCAGGCGGGGCTTGCAGTGAGCCGAGATTGCGCCACTGCACT

63301 TCTTTACAGAAGTTGGGCTGCTCCTGGTGGACAGTGTGTAACAGTGAACAATGTATGCTC
TAGACTGGGTTCCCTTCCTCCACCCTGTGTCTGTGTGGCCTTGGGCAAGTTGTTTAACCA
ACCACTTTTTGCCTCAGTTTCTTTATCGGAACAAGGAGAATAAGAATACTTCAATCAGGC
CAGGCGTGGTGACTCACGCCTGTAATCCCAGCCTTTGGGAGGCCGAGGTGGGTGGATCAC
CTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTCCTTTAC
[T,G]
TATGCTGGCCTGATATTGATCGTCCATGGTAGAATTGATACTGCTTGACAAAGCAGCCTA
TTTCAGTCAGGACCCTTCTTCTCTAGTTTCCTCTGTAGCTATTACCTTAGCCTTCCATTT
CATTCTTCACACTACAGATACTCTCATTGATAAAGGAATGATGTCTTTATGCTTTCAAGC
ATTCTGGCAAGTTAGTAATTCAACTATGATTCTAGGTCAGACAAAACCAGTTATGAACAT
AAGACTGTTTTTAATCTCCTCCCTGGTCCCCCAACCACCCACCCCAATCAGGAGAAACTA

64573 CCTCACCACCCATTACTTTGTGTGACTTTGAGCAAGCTTTAAAACGTCAGTGTCTCAGTT
TTGTCAACTGAGTAGATACCTCATAGAATTGCTGTTGATATTAAGTGACTTAATCCTATG
GGCTGAATTTGTCTCCCAAAGTTCATTTGTTGGAAACTTAATCCCCAGTGCATGTGTTAA
GAAGTGGGACCTTTCAGAGTTGAATAGGCTATGAGGGCTCTGCCTTCATGAATGGATTAA

FIGURE 3-49

TGCCGTTGTTGCAGTAGTGGGTTCCTTATAAAAGGAAGTGTTTGACCCCCTTTCCTTGCC
[T,A]
CCTCATGCATGTGATGGCCTTAGCCATGTTATAATGCAGTAGTAACGCCCTTACCAGACA
CTGGCTCCTTGATCTTGGACTTCTCAGCCTCCAGAACTGTAAGAAATAAAAGTTTTTTCT
TTATAAATTACCCAGTCTCTGGTATTCTGTTATGGCAGCAAAAAACAGACTGAGACACTT
AATATATATGAAGCATCTAGACTGTCTGGCACATTGTACATTTTAAATCCCAGATATCGA
TATCATCAATATCATCATCATCATCTGTGGCTGTATAATACCTCCCTCTGCATTTAA

76462 GCTCGGCTTTGGTCAGCTTCTTTGGTCTTATTTTCCCAAACAAAGAAACCTCTGGGTTAC
GGGCACCCTGTTTACTCCTATCACCTGGCAGGATTTGCAGGATAATTGCCCAGAACTAGA
ATATTGATCCAGATTTTTACATCACCCATCCCTTTTGTTTCTTCTGAGCTGCAGCTGATG
ATCACTGGTTGGTTCACAGAAATAAGCAGGGTTAGTCTAAAATGCAGACAAAAACTTAAA
AACAACTAATGAGACTAGAATTTAATGAAAAGTGTATGATAAATTTTGAAACATAATTTT
[T,C]
CTCTCTCCAGTCCTCATTTTTGTTAAAAACAAATCATGATAGGACTGAGTCATTTGCAGA
ATAAACTTTAGTCTTATATTTGGCCTGGTTATTTGCATAAAGCACAGCAAGAATAATTAT
TTTTCACACAGGCTTTTAAAATTGGCTTTGATGGAACTCTGTTCCACAAGGAATTTCAGA
TAAGACCTTTTAAAGCTGAGCCCAGCCATGGGTTTGTATCCTCAAATACCTATGAGTTGG
GTAAATTCCTCTCTTCTTGAGGTCCCAAGATAACATGGGGTTCCTGGGCCTATTAGAAAG

77652 TCCCATAGCTGATTATAAACCATCTTTTGAAAAGGATCAAAATAAGACAATTGTCTGTGA
ATGACAAAATGTCTTTGGGTAATAACAGTCAAAGCCATGATTGACAAAGAAATTTGGTTA
TTTCTGAGCTTTACAATAACAACATAATAATTTTTTTTTTTTTTTTTTTTTTTTGAGACGG
AGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGCTCACTGCAAGCTC
CGCCTCCCGGGTTCACACCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGC
[G,A]
CCCGCCACTACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTTTT
AGCCGGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGC
TGGGATTACAGGCGTGAGCCACCGCGCCCGGCAACATAATAATTTTAATTACGATTGATA
GCATATACTCAGACATTAGAATTTTAGAAACCTCATAGAATTTTGGAACATATGTATTTT
TCATTAAAATATAACCTGAAGAAGATTAAACATTATTTTTATTTTGGCAATCCCACATAA

77819 TTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGC
TCACTGCAAGCTCCGCCTCCCGGGTTCACACCATTCTCCTGCCTCAGCCTCCCAAGTAGC
TGGGACTACAGGCGCCCGCCACTACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGG
GGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCG
GCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCAACATAATAATTTT
[G,A,C,T]
ATTACGATTGATAGCATATACTCAGACATTAGAATTTTAGAAACCTCATAGAATTTTGGA
ACATATGTATTTTTCATTAAAATATAACCTGAAGAAGATTAAACATTATTTTTATTTTGG
CAATCCCACATAACTAAACATGTCAGTTAATCCTGTTTACCTCTCTTTTGGATGCTCCAG
GAGCCCTCTGTAGTATTCAAAAGTAAGGGGTCAGAAAAGACAACCTTGAAACTGAAGTTT
GATTTTGGGAAGCCTGTTAAGTACATTAGAGGTTTAAAACACTTTATATTATGAAATACA

79594 TAGCTTTTATTTTTCCTTCAGAAAATATTTGATCTAAGTGCTTATTTTTCTCTAAGCCAA
TTAATTAGAGCTCTTTTTTATACAAACATCACACATATTGCACATATATAACTACACAGA
CAGAGGATCCAGTAGTTGTAAGATTTTTCATTGTCCAATCTCCTAATTAGATTACTGACC
TCAGGATGGAGCCCTTCAAGAGCAGGGCTAGGAAAGCATGCAGTTTCTAGGGCCTAATAA
ATAGTTATAGCTGGAAGACAAAAACAGATTTTGAGAGGGATTTATCTGCTTTTAATTCTT
[T,C]
GGGTTTCATGAGGAAAACAGAGGTTTTTTTCTAAAATGGGGTCAGTGGTGCCTCTTCCAT
TTTTTCCAGGGAGTCCCAGGCCATCAGAAGTTATCTTAGGGCCTCTCATGCGTGCATTAA
GAGAGGCAAGACAAAATGGAGAAAAGTAATTCAGTTGACTGAAAAAGAAAATCTTTTTCC
AGTGAAACAAGATGCAAGAAGAGGAAAACATAGAGGCCTTTTAAATATGCCTATAGCTTG
GATATCCACTTTTAATTAAGCTGACTTTTTACCATAGTGCTCTTATTTTAAAAAATCCTT

84331 ATTGACCAAAAGAATCAAATGGGCCTGTTGAAGTGTTCATCTAGTGTCAAGGGAAAATTT
TTCCCCACTGAATAAATTTTAAGAAGGCAGTCAAGACAAGAAGCTATATTTGATTATATC
CTGTTAGTGCTTATTCAATAGACACATAAATCTGTAATTTTTAATATTTGGTATAGAAGT
AGGTTGAAATCCACAGTAATTCACAGAAACTTGTGCAAGGGTTTTGTTTTCTTTTCTTTT
CTTTCTTTTTTTTTTTTTTTTTTGAGACAGAATCTCACTCTGTCCCCCAGGTTGGAGTAC
[A,T]
GTAGGATGACCTCGGCTCACTGCAGCCTCCACCTGCCAAGGTTCAAGCAATTCTTGTGCC
TCAGCCTCCTGAGTAGCTGGGATTACAGGCATGAGCCAACACGGGCGGCTAATTTTTGTA
TTTTTAGTAGAGACGAGGTGTCTCCATGTTGGCCAGGCTGGTCTTGATCCTGACCTCAGG
TGATCTGCCTGCCTTGATCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATGCCCG
GCCAAGGTCTTTTTTCTTGAAAATATCTTCACTCATATAAGCAGTATATGCAATATAAGG

FIGURE 3-50

86107 TCAAGACCAGCTTGAGTAGCAAAGTGAGACCCTGTCTGTACAAAAGAAACACACACAAAA
GAAATATGACTGACTAAAATACATATAATTTTCATAATACTTTAAAATGTAAGAAGGCAA
AAAATTTCTGGGCTCAAGGTGGGTGATCGCTTGAACCTAGGAGTTCAAGACCAGCCTGGG
CAACCTGGCAAAACCTTGTTTCTACAAAAAGTACAAAAATTAGCCAGGCATGGTGGTGCA
CACCTGTGGTTCTAGCTACTTGGAAGATTGAGGTGGGAAATTTGCTTGAGCCTGGGCTGT
[C,T]
GAGATCACAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGCAGCGGAGTGAGACC
TTTTCTCAAAAAAAAAAAAAAAAAAAGGCAAAAAATTAAATTATTAGTATGGTAAAGTTT
CGTTTGGACTTAATATGAAACTCATTTCTAGAAATGATGATCATTTGCATAGGGCTTAAC
TTCCTTTGCTAAGAAAATAGAGTAGTATACTAGGAGACTTCCAGAGCTGCATAGAGCTTC
AGGGTCATCTACCAAGACAGACAATTTGTTGTCATCATCAGTGTTAAACTCTAAATTATT

86175 ACTGACTAAAATACATATAATTTTCATAATACTTTAAAATGTAAGAAGGCAAAAAATTTC
TGGGCTCAAGGTGGGTGATCGCTTGAACCTAGGAGTTCAAGACCAGCCTGGGCAACCTGG
CAAAACCTTGTTTCTACAAAAAGTACAAAAATTAGCCAGGCATGGTGGTGCACACCTGTG
GTTCTAGCTACTTGGAAGATTGAGGTGGGAAATTTGCTTGAGCCTGGGCTGTCGAGATCA
CAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGCAGCGGAGTGAGACCTTTTCTC
[A,-]
AAAAAAAAAAAAAAAAAAGGCAAAAAATTAAATTATTAGTATGGTAAAGTTTCGTTTGGA
CTTAATATGAAACTCATTTCTAGAAATGATGATCATTTGCATAGGGCTTAACTTCCTTTG
CTAAGAAAATAGAGTAGTATACTAGGAGACTTCCAGAGCTGCATAGAGCTTCAGGGTCAT
CTACCAAGACAGACAATTTGTTGTCATCATCAGTGTTAAACTCTAAATTATTAAGTGCTT
ATGTGCCAGATACTGAAGTTTATATACACTTTCTCTAATCTTTAATAATTCTAGAAAGGT

87109 AAAAGATTAAACATATCTATGGTTTTATAAATGATTATAAAATAAATACCCAGTAACTAT
TATCCAGGTCAGCAAATATTCTACTAGTGTATGAGTCAATTTCCATGGCAAAAGAACTAA
GCTTAGGCACTATACTCAAAAAAACTAAAAATAAAATTTTTCTAAATGTGTATTATATCA
ATGGAATAAATACAAATATAACTTACCATGTCATAATTCCCCCCACGCTTTCCCTTCTTT
TACAGCATGGGTAGGTTCTCTCTCCATGGGGATGATTTTCTTTTGCTGCCCAATAGTCAG
[C,T]
GTCTTCACAGACCTATTTGGTTGTCGGAAAACAGCTGTCGTGGGTGCTGCTGTTGGATTT
GTTGGGCTCATGTCCAGTTCTTTTGTAAGGTAAGGACTTGGTTTTTTCATGTTGCTTTTT
AAAAACTGTTAGATACCTTAAAGTTTTACTTTCAGAAACTATGCTATTTACAAGCAAAGA
TCCTCCTTTTCATTTTTTAAAACTTTAAGCAATATGACTTATAAAACAAACTGTTATCCA
TAGCAGCAAATCAGAGCTTGAGAATTTGAATGCTTTTTTTTCTTGTAATGCCTAAGACTT

89444 ACTTATGTAATCTACTACCTAGTTTGTTAACAAAACACACATACAAAGCAATGTTTTTCA
AATTTTTCTGACCACTGAGCAATAAAAATTATGACATATATTTTGATGTGACCCAGTTCT
GTCTCTCTTTCTCTACCCTCTAAGTGAAACAAAATTTATTGAAACCAAAATTCCCTTACT
ACATGTAATATTCTCATATATTCTATTAAATTTCGTTATTTAGCTTGCTGATCAAAGGCT
ACTGAAACTTGAGAGCAAGATACAGGAGCAAGGGGAAATGTGGTATAGATTCTGAGTGTC
[A,T]
AGTGGCAGGTCCATTTTTTCCTCTAGCTCCAGTTCTGCCTTCTGAGGAAAACCTTCTCCA
ACAACTTAGGTCAATCACACCCATGTCCCTTCTCTGAATCCTTTTTGCACATATGATTGG
TATCCGACAGCCTTACTCATTTACATTGCACTTATTTGGCTGCCAAACGTCACAAACTGG
AACCATGTGTTACTGAAGGGAAAACCTGGAAGTGAAAAGGGTTCAGCAGTAGTGCAAATA
CCATCATAAAGCTCATATACTTCACTCTGCAGGAGGGAGAAGCTCTGTGGTTTTCCAACT

90535 TTTTTTTTTTTTTAAATAGAGATGGGGTTTTGCTATGTTGCCAGACTGGTCTCAAGCCAT
CCTCCTGCCTTGGCCACCCAAAGTGTTGGGATTACAGGTGTGAGCCACCACGTCTGGCCA
AGGACCAGATTTTTAATATTCTTTTCCACAATGTATCTGGTACACAGTAGTTGCTTAATA
TGTTGGCTAAACAAAGAGTGGAGATTCAGTAAAGGGTGATCAGAGTGAGGTGAGATTAAT
TTGGGAAAGCCTAGAAGTGATTCTTGAGCCTGATTTGAAGGTGGTGCTAGCTGTGGATTA
[G,A,T,C]
TAGAGGGAGAAGGGCATCTCAGAGAGAGGATTGCCAACATGCCTTAATTTTATCAGATTC
TAGAGTTCCTTATGATTACCTCAGCATGTTGCTAGACTAGCATTATTATCCAAAATTTTA
ATTATTAACCAACTTTAATCTTACTTTCTAACAAATTGTTTGCTTTTACTACTGATAGCC
TTTTCAAAAAACTTTAACTAGTTTTATTCCTTACCATAATTGTTTCAAAGAACATAATGA
TATGATCCTTTATCTTCCTAAGAAATGTGCAATTATTTGGTTAAACTGTAAGATTATTTA

91163 GCCTTACAGCTTACAAACTGGGATCACTAAAGGAATACACTTAATTTAAGTCTTTCTGTA
GTCAGAATATGATTTCTTGTTGTCTTGCACAATACTGAGAACAGTGCAGTACAGGGCGAA
GGTTGGTCTACAGCCCTTAGGCCAGCAAAAACAGGCACAACTGCACCTCTGTGCAAATGT
TCCTGACATAACCTTGGGGAAAAAATATAAAATGCGGCCTTTTCTTTTACTACCTTGTTT
GGTAAGTACCTGGAAAAACTCCATGAAATAATTAGATTTCATAGTTAATTCTAACTTTTT
[T,A]
AAAAAATGTTTCATTGAGACTAGGTTTTTGGTTTGTTAATTGAATCACTGTTGATTTTAC
CCTTCCTGGCACCAACCTTTATTTCTGAGCTGTGGAGAGCACAGTTCTCACTCAGTGCTG

FIGURE 3-51

TGTGCGTCACCTGAAATCCACAGAAAGAGGTGGCTGAACAAAATCACTGATGACCTTAAT
GGTTATTTTTCACATATTCAGATTAAATTAAAATACGTTTAGTGCTACATGCTTGACTTA
CTGAGTTTTTCCCTCTATTTTGGTTAATTTTTTTTTTTTTTGGTTAACTTTTACTTGTAG

93488 CTACATCCTGATCTGATAGTCCATTTTCATACTATTAGGAAAGTATAGCCGAACCAACTT
AAGGTAAGTTTCCTGGAATATAGATCTGTTGTGACAGGATTAACTTTACCATCCAACCTC
TTTCATAGCTTCTGTAGTCAAGAGAACATTTATTGTGTCCTTTCTTAAAAAGATGAGTAG
AAATTCTTTTTCTTTTTTTTCTTTTTTCCAGACAGGGTCTTGTTAAGTTGCTCAGGCTGG
CTTCAAGCAACCCTCCTGCCTCAGCTAGGATTACAGGTGCAAGCCACCACACCCAGCTTT
[A,-]
AAAAAAAAATTCTCTTTGGTACTACCACATGAACACACCTAGAGAAATCATAACTCAGCT
TTGCTAATACTAGACATTTACCAAAGGAAAAGTGGTAGATGACTGTCTAGTTATTTTTGG
TTATATATTTATAATTTGTAAATTAATTTCACATATATTACTTCATTTGACTTTCACAAT
AAACCAGTAAAGCAGATAAAATAAATATTAGCTCCAATTTTACAGACTGAAAAACAGATC
TATTGTTAATAGAGACGTTAAGTGATTTTCCAAGAATTACATGTCAGTAAACAGCAGAGC

96065 GCATAAAAATGGGAGTTATTTTAATGTAAGGCAATGTGATTGCCAACTTGAGATAGAAGT
AAATTTTGAAAGGAGAAAGATAATACCCATTTGGAAAAGTGGTTTTAAAAAGTTTCATAG
CATTGGAGTTGGGCCTTGAGCATGAGATTTTGTGTACAAATCTGATCTTTGATCAACTAG
GGAACTAACTTACCAGTTTAGGTCTTTGAAGATTCAGAAATACAATGGAGTGCTCTCATT
GCTATGTTAAAAATTCTAAGATCTTATTAGATTGTACATGATGATTTGAGAGAGAATATG
[T,C]
ATGCTTGCTTTCAAAGTGAGGTTGGAGGTTTGATCTTCTCGTAGTTGACGTTTCAAAAAG
AAGAATTAGATTGCCTCCTCGAAGCTAAATTTACCTTTCTTTTAGGCCTTCCCACTTAAA
ATCTTTTTTAGAAGGATACAAATCTTATAGATCAATTTAGATGAGGCCTAACTTTCTAAA
AACGATTCCTAGTAGCAGCTGCATCAGTTTTTATGAATTGCCCCTTTTGCCTGAGAGTTG
TTTTGTTTTGTTTTCTGGAATCTTTTTTTGTTTTGTTTTGTTTTGCTTTGTTTTTGTTTT

96351 TGAGAGAGAATATGTATGCTTGCTTTCAAAGTGAGGTTGGAGGTTTGATCTTCTCGTAGT
TGACGTTTCAAAAAGAAGAATTAGATTGCCTCCTCGAAGCTAAATTTACCTTTCTTTTAG
GCCTTCCCACTTAAAATCTTTTTTAGAAGGATACAAATCTTATAGATCAATTTAGATGAG
GCCTAACTTTCTAAAAACGATTCCTAGTAGCAGCTGCATCAGTTTTTATGAATTGCCCCT
TTTGCCTGAGAGTTGTTTTGTTTTGTTTTCTGGAATCTTTTTTTGTTTTGTTTTGTTTTG
[C,T,G,A]
TTTGTTTTTGTTTTCGTTTTTTTTTGAGACGGAGTCTTGCTCTGTCTCCCAGGCTGGAGT
GCAGTGGTGCAATCCCGGCTCACTGCAACCTCTACTTCCCGGATTCAAGTGATTCTCCTG
CCTCAACCTCCCTAGTAGCTGGGATTACAGGCGCCTGCCACCACACCTGACTTAATTTTT
TGTATTTTTAGTAGAGACAGGGTTTTGCCACATTGGCCAGGCTGGTCCCGAACTCCTGAC
CTCAGGTGATCCACCCATCTTGGCCTCCCAAAATGCTGGGATTACGGGTGTGAGCCACCA

96701 CAGGCTGGAGTGCAGTGGTGCAATCCCGGCTCACTGCAACCTCTACTTCCCGGATTCAAG
TGATTCTCCTGCCTCAACCTCCCTAGTAGCTGGGATTACAGGCGCCTGCCACCACACCTG
ACTTAATTTTTTGTATTTTTAGTAGAGACAGGGTTTTGCCACATTGGCCAGGCTGGTCCC
GAACTCCTGACCTCAGGTGATCCACCCATCTTGGCCTCCCAAAATGCTGGGATTACGGGT
GTGAGCCACCACGCCTGGCCTCTGGGTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTT
[T,C,A]
ACGGCTCCTCTGACTCCTCTCATTTAGCTTTCAGGAGCATAAACTCTCTTGGTTTTCTGC
CTACCTCCACATCACTCCTCCTTAGTTTCTTTGCTCACTTCTTCTTTTTCCCACTGACCC
CTGAATATCAGCATGTCCTAGGGCTTGTCCCCTGATCTTTTTCTCCATGTATTCTACTGG
TGGTTTCATCCAGTCTCCTAAGTTCATACATCACGTATATGTCAATGACTTCAAATTTAT
AATTCTGGTCCAGACCTTTTCCCTGAATCCTCCACCAGAGCTGTATATCCAGCTGCTTAC

96879 CCGAACTCCTGACCTCAGGTGATCCACCCATCTTGGCCTCCCAAAATGCTGGGATTACGG
GTGTGAGCCACCACGCCTGGCCTCTGGGTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTCTTT
TTAACGGCTCCTCTGACTCCTCTCATTTAGCTTTCAGGAGCATAAACTCTCTTGGTTTTC
TGCCTACCTCCACATCACTCCTCCTTAGTTTCTTTGCTCACTTCTTCTTTTTCCCACTGA
CCCCTGAATATCAGCATGTCCTAGGGCTTGTCCCCTGATCTTTTTCTCCATGTATTCTAC
[T,-]
GGTGGTTTCATCCAGTCTCCTAAGTTCATACATCACGTATATGTCAATGACTTCAAATTT
ATAATTCTGGTCCAGACCTTTTCCCTGAATCCTCCACCAGAGCTGTATATCCAGCTGCTT
ACTTAACATCTCCACTTGGGTAACTGCTAGGTGTTTCAGACTTACCCTGTCTAACCCTGA
GGTCTTGATCTTACCCCTTAAAACTTACTCTGCCCCCAGCCATCCTCATCTCAGGAGCTG
GCAATTCCGCCCTTTCAGTTGATCAGACTCAAAACTTTGGAGTCATCCTTGGCTCTTCTT

97648 TTCCTTCTCTGCCCCGCGGAGTTTGTTCTCTATGAAGAAGCCACAGGCATTCTTTCTAAA
CATAAGTCACTCTGCTCAGAATCCTTCAATGGCTTCCCATTTCCCTAAGAGTAAAAACCA
ATATCCTTACAGTGACCTACAAGGTCCTTCACAATCTGGCCCCCCACTACCTCTCCGAGCT
TCCATCGCTGTCCCTTGCCCACTCTGCTTCTGCCATTCGCCTTTTAATGGGGCTCACTCT

FIGURE 3-52

GACTACCTGCTTGAAACTTCCTGCGTCCCTTTTCCCCTGAGTATTCACAAACCGCTCCTA
[G,T]
TACTCCTTTTCTTTTTTTGTAGCACTTAATACTTTCTAACATTATCTATTTTACTTCTTT
ATTGTAGTCATTGCTTACTATCCGTATATTTACACGTCTGCTAGAATGTAAACACCACAA
GGGTAAGGATCTATTTCATTCAGTGGTAGATCCCAAGCATCTAGCACAGTGCCTAGCACA
CACTGGGTGCTCAAATATTTGTTGAATGACTAAATATATTCTGGGTGAGTCTGAAGTGAC
ACTGTATAAGTAATGTTCATTTTTTCATCATTTGGATCTTTAAAATTCTCTACTTTGATG

97814 TACCTCTCCGAGCTTCCATCGCTGTCCCTTGCCCACTCTGCTTCTGCCATTCGCCTTTTA
ATGGGGCTCACTCTGACTACCTGCTTGAAACTTCCTGCGTCCCTTTTCCCCTGAGTATTC
ACAAACCGCTCCTAGTACTCCTTTTCTTTTTTTGTAGCACTTAATACTTTCTAACATTAT
CTATTTTACTTCTTTATTGTAGTCATTGCTTACTATCCGTATATTTACACGTCTGCTAGA
ATGTAAACACCACAAGGGTAAGGATCTATTTCATTCAGTGGTAGATCCCAAGCATCTAGC
[A,G]
CAGTGCCTAGCACACACTGGGTGCTCAAATATTTGTTGAATGACTAAATATATTCTGGGT
GAGTCTGAAGTGACACTGTATAAGTAATGTTCATTTTTTCATCATTTGGATCTTTAAAAT
TCTCTACTTTGATGCTATAATGATTTTTCACATTCTGTACTTGCAGGACATGGTGTTATT
AATATTTATTCAATACTTATTCAACAAATAAGCTCAAACTAAGGAAACCTCGGAATAATT
GAGTAACCAGTAATGCTGTCCGTTGATGGAGGAGAGAGTTGGTGTGTTTTGCTCTGATTC

98430 GAAATTTTAAGATAAATAGAAGAAATTTCTGGTCCCTCAAGTAACTGTGTCTTCAGTACC
CACTGAAAAATCTCAAAGAGTCTGGAGTGGTGTGTTTAAGAATAGGATGCAGGATGCAGA
ACCATAACCAGGCCTCAGGTCTGCATAGCTTTGGTCGAGCATTGAGCATAGGGCCTCGTG
AGATAACTGATAAATGCCAAATATGACAATGATAAATGCCAAATATGACAATGATAAATG
CCGAAGAATGACAGTGACAATGATAATGAAGTTACCAAAAATGATGGTAACTTTTCTCAT
[C,T]
GGCATGAAATGCTCTATCTCCAATCTGAAGCTGATGATGTAGTTTCAGTTACTCTCATCT
CTCTCCCCTGCTACTCAGATTGAAAATCAGCTACTTAGTACCTGTGTTCTTTGACTCTAG
ACCATATCATTGGGTCAAATTTCAGTTTTTAAATTTTAGATCCACATGGTTCTCTGTCAA
GAAGATGACTGACTCATATTGAAATCTGTAAAATATGTATTCATTAGCCTGTTTTTTAAA
AACTCCCTTATAAGTGGGTTGACTTTGTGGCAGATAGTAATTGACTGTTCTCAAAAGAAA

101268 CTTGTAGTCACTAACTCTAAGGATCATAGAGCATAAGGGTAAGCAGGCCTTCTTATGTAT
TCATGCTATCAGGAAGGGTCTTTAGCACCCAAACAAAGTTCTAGGGGCTGTACATTGCTG
ATGTGTTAACCCTCAGCTGCCCATGTAGCATCTATTTACCCCTATGCTTTCCCCACTTTC
TATCCCTATCATTATATCTCTGGCTCTTTTGCCCCTCTCTCCTTGGGCAGCTTACTTGTA
ATTAGAAAGTTTATATTCCCTCATAACATATTGTAAAAGTGCTCATTTAAAGGGCAATGC
[A,G]
CACCAAATTGGAGGTGTATAATTGCAAACATGGAATCCCTATATCTCTGTTATGCAATCC
CTGTATCTCTGTATCCATGTTAAATTGAACTGATGCTTTTTTGAAGTAAAATGGTAAGAA
CAGTGGCAACATCTAGTCTTCAGAGCATAGTTTAAGATTTTTGCCCAATCCTCCAACCCA
TGCAATGGTGTGCTTTGAAAACCACAGGTTTCTTTTAGACAAATACAACATTTATTTCCC
GCATTTCTTTTTGATTTAACATTTTAGTTAACATTTTTATTAACATTTTAGTCTACAAGA

103881 ATGGTAAGAGATGGTAAGAGACAACTTTGGCCAGTCAGGGACAACCTCATTGAAAAGCTG
ATAGTAAGTACATCCTTTGGGTAAAGGGTAGTATAAGGTACTTTGAAGGTACAAAAATAA
GACAGCTTTCTATTGCCCTTGGGAGGCCTATAACAGAATTTCTCAAGTCTCTAAGGCCAA
TCAAGAGTTGGATTTTTTTATCCAACTTATTTTTAATTGATGTATTATTAAAAATCTGCA
TATCAAAAATGAAAATGTCTTGCATACTTTGCTGTAGGACCCAATCATTGTTTTTTCTTC
[A,G]
TATACTGCATTAATCTGTTTTCACACTGCTAATAAAGACTTACCTGAGACCAGGTAATTT
AGGAAGAAAAAGAGGTTTAATGGACTTAAAGTTCCACATGGCTGGGTAGGCTTCACAGTC
ATGGTGGAAGATGGAGGAGGATCAAAGGCATGTCTTACATGGTGGCAGGCAGGGGAGTAT
GTGCAGGGGAACTGCCCTTTATAAAACCATCAGATCACATGAGACTTATTCACTGTCACG
AGAATAGCACAAGAAAAACCTGTCCCCATGATTTAATTACCTCCCACCAGCTTGCTCCCA

103926 CTCATTGAAAAGCTGATAGTAAGTACATCCTTTGGGTAAAGGGTAGTATAAGGTACTTTG
AAGGTACAAAAATAAGACAGCTTTCTATTGCCCTTGGGAGGCCTATAACAGAATTTCTCA
AGTCTCTAAGGCCAATCAAGAGTTGGATTTTTTTATCCAACTTATTTTTAATTGATGTAT
TATTAAAAATCTGCATATCAAAAATGAAAATGTCTTGCATACTTTGCTGTAGGACCCAAT
CATTGTTTTTTCTTCATATACTGCATTAATCTGTTTTCACACTGCTAATAAAGACTTACC
[C,T]
GAGACCAGGTAATTTAGGAAGAAAAAGAGGTTTAATGGACTTAAAGTTCCACATGGCTGG
GTAGGCTTCACAGTCATGGTGGAAGATGGAGGAGGATCAAAGGCATGTCTTACATGGTGG
CAGGCAGGGGAGTATGTGCAGGGGAACTGCCCTTTATAAAACCATCAGATCACATGAGAC
TTATTCACTGTCACGAGAATAGCACAAGAAAAACCTGTCCCCATGATTTAATTACCTCCC
ACCAGCTTGCTCCCATGATATGTGGGGATTATGGGAGCTACAATTCAAGATGAAATTTGG

FIGURE 3-53

107845
ACCTTCTCTCCTTCACTTTGTCCCTACTTTTTCTCAGATTTTTTTCAATGATGTCTATGC
TTTTCTTGTTTTTTTCTAGCTTTTCTAACCTTGCATTTATTTTCCTTCAGATCTCAACAT
CTGCTGCAATTTGACTGTTCAGGTTAGTGACAATTTGCCCATTTATCAGTTTTGTGCCTT
AGGCAGTATTCACCAGCATTCTCCTACTTGAGATGAATAAGGATCTTTATTTATCTGACC
ACTTGTTTACTCATTCATGGGGACATTTAATATTTACAGAACACTTTCATCAAAACAAGC
[C,T]
TGTTTTTTCTTTTCAAAATATAATATACTAGCATAGGAACTTGACAGAAGAGGTAATAAT
ACAGAAGAAATCTAGAGAACTGATCATGGAGAAATAATTAAACTAAAACAAAGCTGCTGC
TTATAGTAAGGTAGACCAAGTTTGTCCTGTGTTCCAAATTATACTTAGCCAAAAATAAAT
ATTTATAGATAATTGAATAGTAGTTTTTAGAAATGATTCATGGATTACTCAGGGGTGGAA
ATTATCCCTGTAATGTAGGCCCCAAACTTCTAAAATATTTATAATTTGTGAGGGAGAAAT

109010
AGGTTTAAACCTTAAATAATAGAAATAAAAGTGATTTTATAATTATCTAGAGTAGTTTCA
ATGTGAAATAACTTAAAGGTATGGAAATGGATGCCAAGAAGTATAGTCAGTCTTGCTGGA
GTAAAATAATGCCCAGTGCTTTGTGCCTTCTCCCAGCTGCTGCTTCCAGAAGAACGGGGT
GTCTGAGTGTGAACATCACCCAACAAGTAGGTTAACAGATATCCACGCCCCTCTTGACCC
ACATACATATCAGTGGGATTTAGAATGCTGCCACATATTGATGATTGAATTTATGAAGCA
[-,T]
ATAATATCCTCAATAATAAACCAAGTGTCCCTGTCCCAACTTGTTATCTCTGCTTCTGTG
AACACATGTTTTCTTTTATATGCTCCTTACTCCTCAGGTGCTCTCTCAGGGACTTTTCAG
TTCCTGACCTTGTCCTTTTCAGCATTTTCTCAGAGGACAATTCTTAGCTTCCTGTTGATT
CCTCAAGCATTAATTGCTTTTTTCTGCCAGATATTTCCTTGCTAGGCTCTTGAGCCCTCA
GAGCTGTTCTGAATTATGCAGTGGGAATTGCCCAGGATTAGGAATCACCTAATGTCCCCA

109623
CCTTGTGAGGCACTTCTCAACTCTGCATCCCTTATACCTTCACAGCAACCCTGTGTACCC
AAAGCAGTGTCATACTCGGTGCCTCCTTTTCTCTCCTGAAATAAAATTCCTAGATAAGAA
GACCTCTATATTCCAGGCCTTGTCTTTGATTTTAGGGAAAAAAAAGAAAACTACCTATAT
ACATAATGTTTTTAAAAATCAGTAATGTCCCACTCGTTACAGAAAGGAGAAATAAAAGAA
GTAAGTTAATGCCTGGGATACGTGCTACAACATGGATGAACCATGAGGACATTACACCAA
[G,A,C,T]
TGAAATACTCCAGGCACAAAAGCACGAATACTGTATGGTTCCGCTTAGATGAGGTACCCA
GAGAAGTCACATTCATAAATACTGAAAGTTGTATGGTGGTTTCCAAGGGGAGGGGGAAAT
GAGGAGTTATTTAATGGGCACAGAGTTTCAGCTTGAGAGGAGGTGGTGACAGTTGTACAA
CAATGTAAATGTACTTAATATAGTACACTTAAAATGGTTAAAATGGCAAATTTTATGAAA
TAGGAATTTATCACGATAAAAAATTAAAAAGTAAGAAAAGTTACTGCTTGGGCGAAAGTA

110188
TAAAAAGTAAGAAAAGTTACTGCTTGGGCGAAAGTATATCAAAAAAATAAAAATAGTCCC
CACAAATTTCCAAAACAACCCTAATGAGGTGTTGCTGCCTAAATGGTGAACCAAATTGTG
AACCAATGTGTAGTGTTTGAGACTGGGAAACTGATGCCCAAGATTTTAGCCTCAATAAGG
AGTAGAGTTCATAATTTGACTCCAAAGACATTTCTTTCCCTACCATGCCAAGGCCATCTG
ATTCCCAGTCCAAAGAAGTTTTCTCTCTGCTCTGTAGGCTGCCTTAATCCAGAGTACACA
[A,T,C,G]
GCCTTCCATTTTCTTATCTGTCCTCCTACCAGGGTGTGGTCCTTTTCCTCCTGAACACTG
ACTGTATAATTACCAGACAAAACTAAACATATTTAAAATATAGGCAGTCCTCTACATCCA
AGGTTCCACATCCTTGGATTCAACCAACCATGGATTGAAAATATTTGGGGGAAAAAAAAA
CAATAAAAAAACACTGGCCTGGGCAGCATAGTGAGATGCCATCTCTACAAAAACATTAAA
ATATTAGCTGAGCATTCCAGCACTTTGGGAGGCTGAGGCAGGCAGATCACCTGAGGTCAG

111006
AAAAATAAAAATAAATAAAAATTAGCTGAGCATAGTGGCATGTGCCCATGGTCCCAGCTA
CTTAGGGGGTTGAGGTGGCAGTGAGCTGTGATCGTGCCACTGCACTCCAGCCTAGGCAAC
AGCGAGACCCCATCTCAAAACAAAAACAATAAAACAGAACACAGATTAAAAACAAAATAC
AGGCTGGGCTCACTGGCTTATGCCTGTAATCCCAGAACTTTGAGAGGCCAAGGTGGGAGG
ATTGCTTGTGCTCAGGAGTTTTAGATCAGCCTGGGTAACACGGCAAGACCACATCTCTAC
[C,T,A]
AACAACAACAACAGGAGACTATACTTTCAGGGACCATTTCTGGGGATCATAGTTTGTTAC
TAGAGAAGTTTCTCTGTGTAGAGCATTGAAATATAAAAATGCAGAATAATCATTTACATA
GCATTTACATTGTATTGGTTATTATAAGTAATCTAGAGATTAATTAAAGTATACAGGAGG
ATATACATAGGTTACATGCAAATACTACACCATTTTATATAGGGGACTTGATCATCCATA
GATACGGGTATCTGAGGAGGTGTTGGGTTCAGTTCTCCACGGATACCAAGAGACTAATGT

111223
CTTTGAGAGGCCAAGGTGGGAGGATTGCTTGTGCTCAGGAGTTTTAGATCAGCCTGGGTA
ACACGGCAAGACCACATCTCTACAAACAACAACAACAGGAGACTATACTTTCAGGGACCA
TTTCTGGGGATCATAGTTTGTTACTAGAGAAGTTTCTCTGTGTAGAGCATTGAAATATAA
AAATGCAGAATAATCATTTACATAGCATTTACATTGTATTGGTTATTATAAGTAATCTAG
AGATTAATTAAAGTATACAGGAGGATATACATAGGTTACATGCAAATACTACACCATTTT
[A,G]
TATAGGGGACTTGATCATCCATAGATACGGGTATCTGAGGAGGTGTTGGGTTCAGTTCTC
CACGGATACCAAGAGACTAATGTTAATTTCATTTCCCCAACCTCCACACCAGAACTCTGA

FIGURE 3-54

AATAAGAATAAGAAAAGGAGCAGTTGGGATAGACAATATCAGAAGTATGTGGAAATGATA
ACAGTGGAAGGAAAGCTGATCTAGGCCTACTCAACAAATTTTAATCTTCATTCTGGTAAA
AACAAATTAGATTTATGGGTGCAAATTTGAGCCAGCAATTAGATGGCTCTTAGGATTAAT

111457 ATCTAGAGATTAATTAAAGTATACAGGAGGATATACATAGGTTACATGCAAATACTACAC
CATTTTATATAGGGGACTTGATCATCCATAGATACGGGTATCTGAGGAGGTGTTGGGTTC
AGTTCTCCACGGATACCAAGAGACTAATGTTAATTTCATTTCCCCAACCTCCACACCAGA
ACTCTGAAATAAGAATAAGAAAAGGAGCAGTTGGGATAGACAATATCAGAAGTATGTGGA
AATGATAACAGTGGAAGGAAAGCTGATCTAGGCCTACTCAACAAATTTTAATCTTCATTC
[T,C]
GGTAAAAACAAATTAGATTTATGGGTGCAAATTTGAGCCAGCAATTAGATGGCTCTTAGG
ATTAATAAAAAAAGACTGAACATCATGCCTTCCAAAGACTGAGGGAAAGAGATAGATAGG
AGACTTTGGCAAAGTAGCACTTTAGCCAACATCATTAGCCTAAATCTTAGTGAAGAGAGG
TTAGAAGAAAGGTAGAATTTTCATGGAAGGATCCATTTTTCTTCACTTCAGAATTAAGGG
AAAAATTAGGAAGCTGAATAAGAACTAATGGCCTAATTTCTTTGTTTCTTTCAAAAATCA

112168 CAAGTTACCTCAGATCCCAGTGATTTAAATAATGCTTTCTGAATGTATCCTTTTCTGTTT
TAAGAAGAAGCTGTATTAGGTTCTTGTGTTCCTATAAAGAAATACCTGAGGCCGGATGAT
TTATAAAGAAAAGAGGTTTAATTGGCTCACGATTCTGCAGGCTGTATAGGAAGCATGGCC
CCAGCATCTGCTCAGTTTTCTGGTGAGGTCTCGGGGAGCTTTTACTCATGGCGAAAGCAG
AGTGGAAGCAGCAGGTCACTTGATGAAATTGAGAGCAAGAGTATGGGTGGGGAGCTGCCA
[T,C]
ACTCTTAACCCAATCTCTAGTGAACACAAGCAATAACTCACTTATCACCAAGGGAATGGT
GCTAAGCCACTTGTGATGGATCCACCTCCAAAATCCAGTCACCTCCCACCAGGTCCCACC
TCCAACATTGGGAATCACATTTCAACATGAGATATGGAAGGGACAAACATTCAAACCATA
TCAGAAGCCTATCTTAGGCTGGGCACGGTGGCTCACGCCTGTAATCGCAGCACTTTGAGA
GGCCGAGGCAGGCAGATCATTTGAGGTCAGGAGTTTGAGACCAGCACGGGCAACATGGTG

112653 AAGCCTATCTTAGGCTGGGCACGGTGGCTCACGCCTGTAATCGCAGCACTTTGAGAGGCC
GAGGCAGGCAGATCATTTGAGGTCAGGAGTTTGAGACCAGCACGGGCAACATGGTGAAAC
CCCATCTCTACTAAAAATACAAAAACTAGCTGGGCATGGTGGCACACACCTGTAATCTCA
GCTACTCGGGAGGCTGAGGCAGGAAGCTCTCTTGAACCCGTGGGCAGAGGTTGCAGTGAA
CTGAGATTCTGCCACTGCACTCCAGTCTGGGCAACCGAGTGAGGCTCTGTCTAAAAAAAA
[G,-]
AGAAGCCTATATTAAACTTATAAAATTTAATATCATTTCAACTAGCCTTTTGTTGGGTGC
ATTTGTTCACTTTGGACTATTTTTCCAAATTCATGTACAGTCGTGCATCTCTTAACAATG
AGGATATGTTCTGAGAAATGCATCCTTAGGCAATGTCATTGTTGTGCAAACATCATAGAG
TGTACTTAGACAACCCTACATGGTGTAGTCACTACATACCTAGGCTATATGGCATAGGTA
GAGCCTATTGCTCCTAGGCTACAAACCTGTACAGCATGTTACTGCACTGAATGCTGTAGG

114155 TAAAAATTTAAATAAACTAACAGATTAAACCTTTCATTGTTCAGGAAGGAACAGAACAAT
TTTTGATAACTTGTGAAATATCTGGCACAGAAATTATTTAGAGCCACTAAATAATTTCAA
ATTACCTAAAAATCCTAGTGATTTATTTTCTATTTTAAGATGAAGTCTACTTTAAACTTC
TAAAATGCAGGGTTATTTAAACTGGCATCTAAATCCAAGCTGGTTTTGGTTGGTAATTCC
TCTAGGACATTTTACTAAATCTTGATCTTATCTAAATGATGCTATGTCATAGATGGACTG
[-,A,T]
TTGTTTTGTTTGTAATCCAGGGAAATTAAAAAAAAAAAAAAAACAAGTAGAAATAAAGGCT
TTTAAAGAATTTTTAGAGTTAGAAATGTTTTCAAAATTAGGTTCTTTAAACCATTAGCCA
TCTCTTCCTTCTGAACTCTTCTTTTTTCTGCCCTTTGGTAGCTATGAAATAATCTGCATT
CCAGAAACTTCTTTTTTCCCAGTCCTTTTTCATGTCTTAACAGTGCCATGCATGATTATC
TACACCATGGAAACCCATCTTAATGAAATGGAAAGATCTGCTGTTTAAAAAAACAAACAA

114181 AAACCTTTCATTGTTCAGGAAGGAACAGAACAATTTTTGATAACTTGTGAAATATCTGGC
ACAGAAATTATTTAGAGCCACTAAATAATTTCAAATTACCTAAAAATCCTAGTGATTTAT
TTTCTATTTTAAGATGAAGTCTACTTTAAACTTCTAAAATGCAGGGTTATTTAAACTGGC
ATCTAAATCCAAGCTGGTTTTGGTTGGTAATTCCTCTAGGACATTTTACTAAATCTTGAT
CTTATCTAAATGATGCTATGTCATAGATGGACTGTTTGTTTTGTTTGTAATCCAGGGAAA
[-,T]
TAAAAAAAAAAAAAAACAAGTAGAAATAAAGGCTTTTAAAGAATTTTTAGAGTTAGAAAT
GTTTTCAAAATTAGGTTCTTTAAACCATTAGCCATCTCTTCCTTCTGAACTCTTCTTTTT
TCTGCCCTTTGGTAGCTATGAAATAATCTGCATTCCAGAAACTTCTTTTTTCCCAGTCCT
TTTTCATGTCTTAACAGTGCCATGCATGATTATCTACACCATGGAAACCCATCTTAATGA
AATGGAAAGATCTGCTGTTTAAAAAAACAAACAACCAATCACCCATGCCTCCTACAGTCC

114183 ACCTTTCATTGTTCAGGAAGGAACAGAACAATTTTTGATAACTTGTGAAATATCTGGCAC
AGAAATTATTTAGAGCCACTAAATAATTTCAAATTACCTAAAAATCCTAGTGATTTATTT
TCTATTTTAAGATGAAGTCTACTTTAAACTTCTAAAATGCAGGGTTATTTAAACTGGCAT
CTAAATCCAAGCTGGTTTTGGTTGGTAATTCCTCTAGGACATTTTACTAAATCTTGATCT

FIGURE 3-55

TATCTAAATGATGCTATGTCATAGATGGACTGTTTGTTTTGTTTGTAATCCAGGGAAATT
[A,T]
AAAAAAAAAAAAAACAAGTAGAAATAAAGGCTTTTAAAGAATTTTTAGAGTTAGAAATGT
TTTCAAAATTAGGTTCTTTAAACCATTAGCCATCTCTTCCTTCTGAACTCTTCTTTTTTC
TGCCCTTTGGTAGCTATGAAATAATCTGCATTCCAGAAACTTCTTTTTTCCCAGTCCTTT
TTCATGTCTTAACAGTGCCATGCATGATTATCTACACCATGGAAACCCATCTTAATGAAA
TGGAAAGATCTGCTGTTTAAAAAAACAAACAACCAATCACCCATGCCTCCTACAGTCCTG

115964 GTAGAGGGGTTGTTTTGCTCAGTTCTGCCTAGAAGACAGTTGTAGTTATTTTAGTCCCAC
AGTCTCTACTCCTCCCTGGGCCTGTTTTTGCCCTGCTTTCCTGGCCTCTGTATACGGCCT
TTGTTAGCACTTTGTAGTTGTCAGTCAGCTGTTACTCCAAGTCCTCTGTCAGAAATAAAG
TTACTTTTGTTTGAAGAGTCTCCAGTAATCCCCCTTCTTTTTAAACTATGACTCCCCAAA
GATATATAGTCTAACTTGTGTGCACCAGTATTTTATCATATTTATTTAATAAATTCCATA
[A,C]
ACCTTACATAAAAGATATTAAACAAAAAGTACTCACCCCCTTAAAAGGAGGAAGATAATG
ACCATCAAGGCATGGCCAAATTACACAGTCTTAGGAAACAGTACAGTAGGTAATTATTGG
ACCTTTTTGATTAATGGCATCCTTGCTTTGCAGGTGTGTACCACCACTGCACTTCACAGG
TTTCTGTAAATAAAGTGATTTGGGGTTCCCTTCTGTTGAGCATTCTTCTATCGTGAAACC
ATTTGGTGATGAGAAGCTTGAGTTTCTAATGCATGTTCGTTGGCTTTATTGGAGCTGCTT

118100 GTGGTGGCTCATGCACTTTAGGAAGCCGAAGCAGGCAGATCACTTGAGGTCAGGAGTTTT
GAGACCAGTCTGGCCAACTTGGTGAAACTTCATCTCTACTAAAAATACAAAAATTAGCTG
GGCGTGGTGGTGCACACCTGTAATCCCAGCCACTCTGGAGGCTGAGGGAGGAGAATCGCT
TGAACCAGGTAGGTGGAGGTTGCAGTGAGCTGAGATCACACCACTGCACTCCAGCCTGGG
CGACAGCGAAACCCCATCTCAAATAAATAAATAAATAAATAAATAAATAAATAAAATAAT
[-,A,G]
CCCCTTCTAGTTGAAAAACTAAGTTCCTACCTAAGCATAATTTGGATTTACCCAATTTAT
CTTCTTTCAAAATACCTCAAACATTTCACCTTATTATTCTTTTTAAGGATTACAAAGTAG
AGCAGGGGGGAAATAATAAACCACTAATAAAGAATAATAGCCATTTGACAGACAGGTGTT
CTTAGTTTCATAAAAAAAAAAGATGCCTGGTAGATTCAGTCTTTTATGAATACTAAAGAAT
GCCTCTTATTTTTGTTTGTTGGAGACAGGTTCATTTTGAACCTAACCTGGTGTTCTCAGG

119631 GTAGGTAATAAAGGTACAAAGATAGATTAGCAACAGATTACGGAGAGCTTTGAATACCAA
TCTAAGGAGTGTAAATGTAGGCAGTGGGCCACCTTTGAATAAGGAATTGATGAGATTAAA
GCCATATTTAGGAGGATTATTCTGGACCAGTATGAAAACACAGAAGTTAGGGAAAACAGT
TAATAGTTTTGAAAAGAGAAGAGAAAAAGGAGATGGTGTTGGGATACATAAATGGGCTTTT
AAAATGCAAAATGAGAAGTGTTTTAAAGAGATATCACCCAGAAAGTCTATGCACTGCCAC
[A,G]
TGGGCACTATATGGGTGGTTGTATTTGGTGGGAAATTTGCTTGCAGACTTCCAGAACTCA
GACCAATGTGTGGTGTGGGGGACGGTGATTGTCAGGCATTATGGAAAGGTCAAACAAAAT
ATGCTCACTGGCTATCTATGGCCCACAGGTCACTGTAGTCTCTGTTATAAGTACACTAAG
TGGAGGAGAAAGGTCCTTTAAAAAAAAGAAAGCTAAAATTAATACCTGATTGTTATTAAC
TGTGTGCCAAACACTGTTCTAAGCTCTTTACACAGACATTTTATTTAATCCTCGCAACCA

120833 CAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGTCACCACGCCTGGCTAATTTTTATAT
TTTTAGTAGAGATGAGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGGCCTCAG
GTGATCCACCCACCTCGACCTCCCAAAATGCTGGGATTACAGGCGTGAGCCACCATGCCC
GGCCTTAAAAATGCTTTTAAAAATGAAAACTAAAACATGTTAATTTTTTTCAAATGTTTT
CATGAAAATTATCACAGGACAAGTTTCATAAATATTGAAATTTGGAAAAAGTTGCAAGCC
[T,C]
ATAACATTGCAGAGAAGCAAATGCATTTGATGCAAAGCCTCAAATTTGTCAAGTTTTTCT
ACCATATTCAGTGTGGTTTCTTTCTCTTTGGCCTATAGATGAAACATGTAAATGAAAGAT
TTCAAGATGAAAAAAATAAAGAGGTTGTTCTCATGTGCATTGGCGTCACTTCAGGAGTTG
GACGACTGCTCTTTGGCCGGATTGCAGATTATGTGCCTGGTGTGAAGAAGGTTTATCTAC
AGGTACTTTTTTACACCTTTTTTCCCCTATCAAAAATTACTCTCATCACCCAATGTCTCA

121125 TGCAAGCCTATAACATTGCAGAGAAGCAAATGCATTTGATGCAAAGCCTCAAATTTGTCA
AGTTTTTCTACCATATTCAGTGTGGTTTCTTTCTCTTTGGCCTATAGATGAAACATGTAA
ATGAAAGATTTCAAGATGAAAAAAATAAAGAGGTTGTTCTCATGTGCATTGGCGTCACTT
CAGGAGTTGGACGACTGCTCTTTGGCCGGATTGCAGATTATGTGCCTGGTGTGAAGAAGG
TTTATCTACAGGTACTTTTTTACACCTTTTTTCCCCTATCAAAAATTACTCTCATCACCC
[A,G]
ATGTCTCATTAAATGTACTTACATGCTTAAATTCTTTTTTTTTCTTTCCTTTTTCTTTTT
GAGATGGAGTTTCGCTCTTATTGCCCAGGCTGGAGTGCAATGGCACGATCTCAGCTCTCC
GCAACCTCCACCTCCCGGGTTCAAGGGATTCTCCTGCCTTAGCCTCCCAGGTAGCTGGGA
TTACAGGCGTGTGCCACCACACCAGGCTAATTTTTGTATTTTTTTAGTAGAGATGGGGTT
TCTCCATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGC

FIGURE 3-56

121245
ATGAAAGATTTCAAGATGAAAAAAATAAAGAGGTTGTTCTCATGTGCATTGGCGTCACTT
CAGGAGTTGGACGACTGCTCTTTGGCCGGATTGCAGATTATGTGCCTGGTGTGAAGAAGG
TTTATCTACAGGTACTTTTTTACACCTTTTTTCCCCTATCAAAAATTACTCTCATCACCC
AATGTCTCATTAAATGTACTTACATGCTTAAATTCTTTTTTTTTCTTTCCTTTTTCTTTT
TGAGATGGAGTTTCGCTCTTATTGCCCAGGCTGGAGTGCAATGGCACGATCTCAGCTCTC
[C,T]
GCAACCTCCACCTCCCGGGTTCAAGGGATTCTCCTGCCTTAGCCTCCCAGGTAGCTGGGA
TTACAGGCGTGTGCCACCACACCAGGCTAATTTTTGTATTTTTTTAGTAGAGATGGGGTT
TCTCCATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGC
CTCCCAAAGTGGCTTAAATTCTTCTATAAAAATGAGAAATATTTTCTACAACATAACTTC
TATAGGCAGTTTTTCAAGGACAAAATTAGTTATTAGTTTGGGTTTTAAACATGAGAAATT

121521
TGCAATGGCACGATCTCAGCTCTCCGCAACCTCCACCTCCCGGGTTCAAGGGATTCTCCT
GCCTTAGCCTCCCAGGTAGCTGGGATTACAGGCGTGTGCCACCACACCAGGCTAATTTTT
GTATTTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACTCCTGA
CCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGGCTTAAATTCTTCTATAAAAATGA
GAAATATTTTCTACAACATAACTTCTATAGGCAGTTTTTCAAGGACAAAATTAGTTATTA
[G,A]
TTTGGGTTTTAAACATGAGAAATTGGCAATGAAACAACAATTTCTTTGTTTTGTCGTGGA
ACTCCACCAAACCAGAATGGTTTTCATCCATTGCTTTTTCTATGAAGAATGTTTTTTGGT
GTAGTTCTCATAGTCATGTGCAGATCCTGTGCCCTTTGCATGTCTTATGAAATTTGGTTG
TGTGTGTGACTTTTCAGCTTCTTTACTGCAAATTGCCTCCTCGTGTTTTGGGGTGAGCAT
AAACAAATGCTAATTCCAAGATCATTGCTGACAATCAACAGAACAGGTATTGAAGTGACT

124296
GCATGAAGTAAGTGGAGGACTAAGGAGATGGAAAGGGAGTGGCCAGATAGGCAGGGGGAA
AACAGGGATGAGAGTGTTTCTAGGGGCAGGAGTGGTCAGCAGTGTCAGTAGCTAGTGACA
AAGAGGCTGAAAGGTTTTCATTGAATTTAGAATAGGGAGGCTATGAGTGAGCTTACAGAG
AATGGTTTCTACAGGTGCAGAAAATTGATCATGATCATTTGAACAAATGGAAATATAGAC
AACTTTGTCCAAATGCTTGGCTGTGGAAGGAAGGTAAAGGCAAGCCACAAGGGGGGGTCT
[C,T]
AGGTTTTGAGGATTAAGCCTGTCTAATTATATTATACATGAGAGGCAGTATTTGAGCAGG
GCCTTGAAAAAGAGGTGAAATTTGGACACAGGAAAATGATTGGAAAAGGCATTACAGATA
AAGTTAACTTCCATTAATTGACTTGAAGTAATAACAGTCAACCATTTATTGAGGACTTTC
ATGTGCCAGACAATGAACTAAGGGCTCTACATGCATTATCTCATTGATCCTTGCCCCAGC
CCTTTAAGAGAGAAGGTACCATTGTTATTTCCACTTAGAGATGTGAACACTGAGGAACTG

124549
TGCTTGGCTGTGGAAGGAAGGTAAAGGCAAGCCACAAGGGGGGGTCTTAGGTTTTGAGGA
TTAAGCCTGTCTAATTATATTATACATGAGAGGCAGTATTTGAGCAGGGCCTTGAAAAAG
AGGTGAAATTTGGACACAGGAAAATGATTGGAAAAGGCATTACAGATAAAGTTAACTTCC
ATTAATTGACTTGAAGTAATAACAGTCAACCATTTATTGAGGACTTTCATGTGCCAGACA
ATGAACTAAGGGCTCTACATGCATTATCTCATTGATCCTTGCCCCAGCCCTTTAAGAGAG
[G,A]
AGGTACCATTGTTATTTCCACTTAGAGATGTGAACACTGAGGAACTGAGAGGCTACCTCA
CTGTGGGTGTCTGTGTGTATAGAGGTCCAGGCAGTCACAGGACAGTCTGGTCACACAGCT
AGAAGGAAGTGAAGAGAGTTGGTAATGTGCTGCTTTTAAAAATGTATTTATTGTATCATG
ATCACTTTGTCAGTACTTCACTGTGGACATCCTCATCTAACATTTAGTTTTGTTCTCTAG
TGTCAAGGGAGCCCTCTAATGGACATTTATTGCACTACAGACTTTCAGCTTTCATACATT

124858
TTGTTATTTCCACTTAGAGATGTGAACACTGAGGAACTGAGAGGCTACCTCACTGTGGGT
GTCTGTGTGTATAGAGGTCCAGGCAGTCACAGGACAGTCTGGTCACACAGCTAGAAGGAA
GTGAAGAGAGTTGGTAATGTGCTGCTTTTAAAAATGTATTTATTGTATCATGATCACTTT
GTCAGTACTTCACTGTGGACATCCTCATCTAACATTTAGTTTTGTTCTCTAGTGTCAAGG
GAGCCCTCTAATGGACATTTATTGCACTACAGACTTTCAGCTTTCATACATTCAAAAATT
[G,T]
AGTGCCTCCTGTGCCCCAAGCACCAGCTCAGATGCTGTTAGGGTGATGCAAACAAGACAG
ACATGGTCCCTGAGTTCCTAAAGCAAGCCTGAGGCAGGAAGAAGCCGAATGTGTGTGGAA
ACCCAAGAAGATGGGGAAAGTGGCATGGGAAGGGACTGGAAAGTTAGAGTGGGTCAGATT
AGACAGAGCCTTGAAAGCCAGGATGAAGAGACTTTGCTTTAAGAGTAGTGGATTTTGGCC
GGGCGCAGTGGTTCACGCCTGTAACCCCAGCACTTTGAGAGGCCAAGGCTGGCGGATCAC

125920
TACCCTAGATAGGGCAGACAGATTATCTGCAACATTTTTGGAGCACATTTTAATACCTGA
CTGTTTCCAGTAATTTACAAAAGAAAATATAGCCTTTCCTAACTTGTCCCATGTTGGTCT
GCAGTTACACAGCAGTAAGTTAAAAGTTAGTATTGGGGGTCAAATATTTCACTTTAGATG
AAAGTTTAGCCACAATCTGGCTTCTGTTAGGCCTTATCTAATTTTTGCATCCAAATGTAG
AGCATCGTTTGTGGCACCCAGTAGCACATGCTGAGTCACAGGTGTGACAGCTGCATTTCA
[A,T]
ACAAGCCTGAGAAGGAGAAAGAAAGCCCTTCAGTGTGTCCTGTGGTTGCGAGGAGCCACT
CACGGACTCCACCTTGTGAACACAGCGGCACAGGACGCAACACAGGCCTAACCCATGCAG

FIGURE 3-57

GATGCTGGACTCGTTCCCTTATTCACTACCTCCTCTCTCTCCTTTTTCATGGCTTCCTTG
CCCCAACATCCCAAACACACAGTGTGGTTTTTGATTCTTGGCTCTTCTCTGCTGCAGTTG
ACTCCAGCTCTGCCTGTTTGTTCCTTCCTTTTTCTCCATCCCTGGCTCTCTGCCTTTTGG

126266 TTGCGAGGAGCCACTCACGGACTCCACCTTGTGAACACAGCGGCACAGGACGCAACACAG
GCCTAACCCATGCAGGATGCTGGACTCGTTCCCTTATTCACTACCTCCTCTCTCTCCTTT
TTCATGGCTTCCTTGCCCCAACATCCCAAACACACAGTGTGGTTTTTGATTCTTGGCTCT
TCTCTGCTGCAGTTGACTCCAGCTCTGCCTGTTTGTTCCTTCCTTTTTCTCCATCCCTGG
CTCTCTGCCTTTTGGCCCATCCTTTAAGGCTTGGAATGCTCCTGGGCTTACCTTCTTTCC
[A,G]
TTCATTGGTGGTGTGTTTTCTCAGACATACCTGCTCCCTGCTTTCATCTTTCAACTTCTT
GGGGCTGTGACAACTCTTCCCTCTTTGTCCCCTGGAAGCCAGTTCTGAGTAGCAGCCAGG
CCTAGAACACTGGTGACACAGACACACTTCATAGCCCTCCCGCATGGTCTAGTTTCAGAT
CATGGTAATCCCTAGTCTAGGAGGCTGCCGAGCCCAAGAGCACAGGCTCTGGAGTGAGAA
GCCAGTTCACCCCAGTTCTACCACCAACTTGAGAATCAGCAGAGGGCTGTGGTGAGGATT

128258 TGATCAGAGGCTCCGTACTCAGTCTCATCTAGACTGTGGCACTGGGTGTGAACGTATCAA
ATGCATGTTTCTCCATCAGGCAGAAGTGAGAGTAACCATGTGCCATGGAGAAGGTTGACA
GACTCCCTGTGAAGCACTTCGAAGTGACACTGGCCTCTGTGTGCTTCAGAAGAATCCAGC
CACCTGCTGTGTGGCCTGACATTTTCCTTTAGTTTGTGATGGGCCAGCAGAACTCTGTTG
CCAACTGTTTTCTGTCCTGGGTGCCTAGCCAGAGGTTCTGAAAGTCTGGAGACTTTATAT
[G,T]
GGCTAAACTTTAGGAACGTCAATTACATGTCTATCTCCAAGATGCCTTCTTTTATTCAGG
TGCAGCTCATTGTTTCCTCTTGAGCTACACTTAAGATTCTTGAGCAAAACCTAAACTGAC
ATTTCTCCAGCAATGCTCTCCTTGAGATAGAAATGGGAAAAGTAAGAGCAAAAGGAATCT
TTTGTTCTCATGTGCATACACTAACTCATAGAAGGTTAATACTTCTATAGCCTGTACTAT
TATAACAAGTATTATATATTTATGATATATTTCCTTAAAGAAAACAAAAGCAATATAGAC

130303 TTTATCTCCTTTTATGTTTGTACATAGAATAAAAATGTTTCTATTGTTAAGAATATTAGA
GTTGGACGCAGTGGCTCACGCCTATAATCCCAGCACTTTGGGAGGCCAAAGCAAGTAGTT
TGTTTGAGCCCGGGAGTTCAAGAATGGCCTGGGCAACATAGTAAGACCCCATCTCTACAA
AAAATAAAAAATTAGCCGGGCATAGTGGCATGTCCCAGCTACTTGGGAGACTAAAGTGGG
AGGATCACTTTGAGCCCAGGAGGTTGAGGCTGCAGTGAGCTATGATCGCACCACTGCATT
[C,A]
CAGCCTGGGCAACAAAGTGAGACCTGTTTCAAAAATAAAAAATTGGGGTTTATCTACTTA
GATTTTCAATAAAAATTACTACTTAAATCTTTACCTGCTTGTTAATTTCAAACCCTTTTC
TACATTTTGATTTATCTTTAAATCTCTTTTTGTCTCAATAAATGGGAAGTATCAGGAAGT
CTTTTTACTTGCTCAAGGTCATAGAGAGCTTAGAACCTGGTAGTGTCCCTCTGAGCCCCA
GTTCTTTCCAACCTGCCAGGCTGTAGGCCCAACAATTACTCACCACTAAGAAATTATGCT

130617 AAAGTGAGACCTGTTTCAAAAATAAAAAATTGGGGTTTATCTACTTAGATTTTCAATAAA
AATTACTACTTAAATCTTTACCTGCTTGTTAATTTCAAACCCTTTTCTACATTTTGATTT
ATCTTTAAATCTCTTTTTGTCTCAATAAATGGGAAGTATCAGGAAGTCTTTTTACTTGCT
CAAGGTCATAGAGAGCTTAGAACCTGGTAGTGTCCCTCTGAGCCCCAGTTCTTTCCAACC
TGCCAGGCTGTAGGCCCAACAATTACTCACCACTAAGAAATTATGCTTGTGCTGTCATGG
[C,A]
AGTTGCATTGGAGAAAAGGATATTTAACTGGCAAACAAAAGTCAGGAGAATGGGGAGATT
TTGTTCTTTTGAAATGCTAGTGTGAAGTGCTAGGCTTATTTTTCAAATGCCCAACTCGTA
TTCTTTTCTTTTCTTTTTTTTTTTGAGAGGGAGTCTCACACTGTCGCCCAGGCTGGAGTGC
AGTGGGGCGATCTCGGCTCACTGCAAGCTCCGCCTGCTGGGTTGACGCCATTCTCCTGCC
TCAGCCTCCGAGTAACTGGGACTACAGGCGCCCACCACCACGCCCGGCTAACTTTTTTTT

130910 GTCATGGCAGTTGCATTGGAGAAAAGGATATTTAACTGGCAAACAAAAGTCAGGAGAATG
GGGAGATTTTGTTCTTTTGAAATGCTAGTGTGAAGTGCTAGGCTTATTTTTCAAATGCCC
AACTCGTATTCTTTTCTTTTCTTTTTTTTTTTGAGAGGGAGTCTCACACTGTCGCCCAGGC
TGGAGTGCAGTGGGGCGATCTCGGCTCACTGCAAGCTCCGCCTGCTGGGTTGACGCCATT
CTCCTGCCTCAGCCTCCGAGTAACTGGGACTACAGGCGCCCACCACCACGCCCGGCTAAC
[-,T]
TTTTTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGAT
CTTGATCTCCTGACCTCGTGATCCGCCCTCCTCAGCCTCCCAAACTGCTGGGATTACAGG
CGTGAGCCACCGCGCCCAGCCGGCCAACTCGTATTCCTAAACGAATCATAATTTTACCAT
AAGACCATAGTTTAGTGATTGAAGAAAAAATGTACCGAACTGTATGATATGATGGTGTCA
AAAAGAACTAACCCAATATGAAACAGTTTTCAGGAGCATGTTTCCTATTTTGGTGTCAGT

131727 TCCTCCCTAGCCTGACTGCTATTGGAGGGCACCCTCCAGGCACAGTTCTGTTCACAGCCT
GCTGCTGCCCCCGTGGGCCATGGTTCCAGGACGGCTCCATCTTCTGTGCTTTGGGCACAT
TAACCTCTCCCAGCGTCACTATCTTCATCAGCAAAATGGAGATAACATTAGTACCACCTC
ATAAAGTTGTTATGAGGATCACCAGTGAGATAATCAATCTAAAGTGCTTACAACAATGCT

FIGURE 3-58

TGGCACTTGGTAAACACTAAATAAATGATAGTTGCTATTATATGCATACTTTTAAAAAAC
[C,T]
TGATGCTTTTAAAAATTTTTTCTGCTGACTAGTGAATTGTTCAGTTTTTGTTGTTGTTGT
TGTTGTTGTTGTTGTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGC
ATGATCTCGGCTCACTGCAAGCTCCACCCCCCGGATTCACGCCATTCTCCTGCCTTAGCC
TCCCGAGTAGCTGGGATTACAGGCGCCGGCCACCACGCCCGGCTAATTTTTTTGTATTTT
TAGTAAAGACGGGGTTTCACCTTGTTAGCCAGGATGGTCTCGATCTCCCGACCTCATGAT

132895 CAGGCTGGGAGGGAAATTAAATATATAAAAAGTCATTTTCGGCCAGGCGCTGTGGCTCAT
GCCTGTAATCCCAGCACTTTGGGAGGCCGGGCAGGTGGATCACTTGAGGTCAGGAGTTCA
AGACCAGCCTGCCCAACATGGGGAAGCCCCATCTCTACTAAAAATACAAAAATTAGCTGG
ACTTTGTGGTGCTTGCCTGTAGTCCCAGCTACTCAGGAGGCTTAGGCAGAAGAATTGCTT
GAACCTGGGAAGCGGAGGTTGTAATGAGCTGAGATCACACCACTGCACTCCAGCCTGGGC
[G,A]
ACAGAGACAGACTCCAGCTCAAAAACAATACGTTTTTTTAATCTTGTCCCTTAATGGAAA
TATTGAGAAAATGTCTAGGGGAACTCGAAGGAGATGATTATAGGAGTTGATTATGTATGT
AAAATCAAAGTGAATGAGGCAGTGGCAGGGGGGAAAGGGGGAACCAGTAATGACTTAGAA
GTCCTAAGCATGTTGCATGGTAATTGTGACATTTGCTTCCTGCGAGCGGAGCTGACCTTG
TGGTGTCCGTCCTAGGTACTCTCCTTTTTCTTCATTGGTCTGATGTCCATGATGATTCCT

133506 TCTTTGGGGCCCTCATTGCTGTGTGCCTCATCATGGGTCTCTTCGATGGATGCTTCATTT
CCATTATGGCTCCCATAGCCTTTGAGTTAGTTGGTGCCCAGGATGTCTCCCAAGCAATTG
GATTTCTGCTCGGATTCATGTCTATACCCATGACTGTTGGCCCACCCATTGCAGGTAAAT
ATAATGATTCTCCAGTAGTTATATTAATTCATAGTATTTTCTACTTCAGGTCTTAATTAA
GTCTCATTTATATGTAAAACATATTACAGGTTATTGATTACTGGTCTTTTGCTTTTATGT
[G,A]
TGCCTTACTGGTAACATTTTTAATAAGACTAGCTATTAAACAGTATGTTGAATTTGCTGA
AGAGTCCCTCTTATCCTTACTGGCAGCTAAAACACTATATAAAACACTTTGGGAGGCTGA
GGCGGGAAGATCACCTGAGCTCAGGAGCTTGAGACCAGCCTGGACAACATAGTGAGACAC
TATCTCTACAAAAAATAAAAAATAAAAAAATTAGCCAGGCGTGGTGGCATGCATCTGCAG
TCCCAGATGCACAGGAGGCTGAGGTGGGGGGATTGCTTGAGCCCAGAGGTCAAGGCTGCA

135473 ACACCACCACACAGCAACTGGCAGCAGAAAGCTGCTGGCTCACCAGGAGGCTCTGAGGGC
ATCATAGTATTACAGCAGAACATAATCATGAGTGTTCATAAGGAGGAAGAGTTCAGGAAT
GGTCTGATTCCAAAGATAGAAAACTTGTGCATTTACAGATATTGTCTATTTTATATATGC
AGAATGGATGTACACATTTTTATGTGTTGGTACTTAATTGGACAAATGGTCAGCTATTTA
TTTAACAATGTCAGATAGATATTTAGTACAGTATATTCCCTGGACATTTTAGATAACTTA
[G,A]
GTTTTATAGTATAAGTTATAAGAGTTTAAAAACACTAAGATAACTTTTAAAACCATGAGT
CCTGGAATTTGTTAGAGAAATTAGAAATGTTGAGTACCGTAAAAGTTTTCAGAAGCAGAA
CCCGAATATAGAATGCCATTAATAATTCTTATATACTTACATTTTCTTCCAGGTACTTAT
AAACCTTAGTCTATTTCCAAAATGTATTTAACAACCTAACTTTCATCCAAATTATTAAAT
TTAAAATTTTTCTTAATTTTGAATTATGATGTTAAATAGTTTTCATGTTATTTCATAAAA

136201 GACAAAACCCCATCTCTACAAAAAATATAAAAATTAGCCAGGCATGGTAGCATGCACCTG
TAGTCCCAGCTACTCAAGAGGCTGAGCTGGGAGGATCAATTGAGCTTGGGAAGTCAAGGC
CTGCATGAGCCATGATTATGCTACCGCACTCCAGCCTGGGTGACACAGTGAGACCCTGTC
TCAAAAAAAAAAAAAACGTCTACCTTTCTATCTTTTCAAGTAATGTCTTGGTATGAAAATC
CAGAGGACTGTACTTTATGACAACGTTAAAGATATGAGATCTTTTTCTTCCTATCAAAAA
[A,G]
GGTTTATATTTCAGATCTGTTTCCTAAAAAAAAAAAAAAAGTCTGATGTCTGAGATGTCTG
AATCAGTCCTGTGGCTGATCTAGACCCCCTACAGAGCCACACTTGTTCTCCCTTGGTGAC
AGCTTTTTCCCTTCTCAGGGTTACTTCGTGACAAACTGGGCTCCTATGATGTGGCATTCT
ACCTCGCTGGAGTCCCTCCCCTTATTGGAGGTGCTGTGCTTTGTTTTATCCCGTGGATCC
ATAGTAAGAAGCAAAGAGAGATCAGTAAAACCACTGGAAAAGAAAAGATGGAGAAAATGT

137080 TGGTATATGAAAACATGTCTGAAAGTCACATATTGTGAAAATTTGAAGCTATCTCAGTAA
AAAGCAGCTTTGGAAACTGTGAATGATCTTTAGCTTGTACAAATGTTTAAAAATACCTCA
GGCTATACTGAAAGGGTTGCAGTTTGGTTAGGAGTGGAAATATTTTGTTTGTTAATGATG
TCTTCAGTTCTGGTACCTCTGTTTTACTTTCTTATGCTCTTTGGAAACTTTTTGCAAAAT
TTAAGCCTGGGTTCTAGATAATACCAGATCTACCTAAACCTCAAGTCTATGTTAAAGTTG
[A,C]
TTTCCTGCTGTTAAATAAGCTATGATATTAAGATATTCTGACTTGCTCCAGTGTCAAGGG
ACCTTCTGGGAGCAGGTGCTAACATAGTGTTCAGAATCAATATGTGAGATGAAAAGGATC
CCCTCCAGGAGGATCCTGAGCTGTTCAGAAATCATTTAAGTTTACAGCGTTGTTCCCTTT
GCGTTTGCAGTGCGTTTTACTCAAGTAGCCAGAAACACCCCACGTTTCTGAATTTGTTTA
AACTGTAACAATAAAGTAAAATAGAATGCATGAAAGATATTCTGGCGATTGTAACTTAGA

FIGURE 3-59

138022 GAAAGCTCTCAGTTTGAGGACCCAAAATAAAACCAAAGTCATGCCATGACCCATACTCAT
TTACAAAAACAAGAACACTTTCCTCTATCCCTAAAATTATGCTTTAGTACTTGAGGCCTT
TAAAAGTTAGTGCTTTTGATTGTGAAGACATTCAGCAACTTACTTTGTCATACATGCAGT
TGCACCTTACCACTTCTAATAGTGTCATATTTCATATTCAGGGGACTTAGATAATTTGCC
TGTGGATGGTTCTTTTGCAGGAAAAAAAATCTACATTTTGACCATACTACCCTTTCATGT
[T,C]
CTTATTATAAGCTTTTAGAAAATGATTTCATTCAGTCATGTCCAGTTATATAAAACGTTA
CTTTCTCATTTTTGAGAAGTTCAACAAAACATACTACTAAGACCAATCATCAAACCCACT
ATTATAAATGTTAATTTTGGCTGGGTAAGGTGGCTTGCGCCTATAATCCCAGCACTTTGG
GAGGCTGAGGAGGGAAGATTGCTTGAGCCCAGGAGTTTGAGACCAGCCTGGGCAACATAG
CAAGATCCTGTCTCTACTAAAAATAAAAAAAAAAATTAGGCCAAGCATAGTGGCTCATGC

138543 GACCAGCCTGGGCAACATAGCAAGATCCTGTCTCTACTAAAAATAAAAAAAAAAATTAGG
CCAAGCATAGTGGCTCATGCCTGTAATCCTAGCACTTTGGTAGTCCAAGGCAGGGGGATC
ACTTGAGCCCAGAAGTTCAAGACCAGCTTGGGTAACATAATGAGACCCTGTGTCTACAAA
AAATTTAAAAATTAGCCAGGCATGATGGTGCCCACCTGTAGCCCCAGCTACTCAGAGGCT
GAGGTAGTGGAAGGATTGCTTGAGCCTAAGAGATGGAGGCTGCAGTGAGCTATGCCACTG
[A,T]
ACTCTAGCCTGTTCAACTGAGCAGAACCCTGTCTGTAAAAGAAAATCAAAAACAAAAAAT
AAATGTTAAATTTTGTTTTAAGTTTTAGCACAGACTCCCCTCAAAACACCTTCTCCCCAA
TTTTACAGAAAGTAATTCAAAAATGAAAACTTTACTCTGTAAAGACCTCTACAGTGTTTT
TCTTTTCAAAATTTGGCTGATTTTAGGAAAAAAGTGATCATCTGAAACTAAAAGAAATTG
CTTGGTTAGTTTCCATATTAAAACAGCAGTGACAAGTATATATAACTTAGATCTCAGCAT

138681 AAGACCAGCTTGGGTAACATAATGAGACCCTGTGTCTACAAAAAATTTAAAAATTAGCCA
GGCATGATGGTGCCCACCTGTAGCCCCAGCTACTCAGAGGCTGAGGTAGTGGAAGGATTG
CTTGAGCCTAAGAGATGGAGGCTGCAGTGAGCTATGCCACTGTACTCTAGCCTGTTCAAC
TGAGCAGAACCCTGTCTGTAAAAGAAAATCAAAAACAAAAAATAAATGTTAAATTTTGTT
TTAAGTTTTAGCACAGACTCCCCTCAAAACACCTTCTCCCCAATTTTACAGAAAGTAATT
[C,T,G,A]
AAAAATGAAAACTTTACTCTGTAAAGACCTCTACAGTGTTTTTCTTTTCAAAATTTGGCT
GATTTTAGGAAAAAAGTGATCATCTGAAACTAAAAGAAATTGCTTGGTTAGTTTCCATAT
TAAAACAGCAGTGACAAGTATATATAACTTAGATCTCAGCATATGTGTTTGTATATTAAA
CTTCACATATGTAGTTTTCAGTTTAATGGAATGAATCAAACTGGATCTATAACACTGAAA
AAGTTCTATTGTAATAGACTCATACGGAGAATACTCTGCTATAATAATATAAAATTAAGA

FIGURE 3-60

SEQUENCE LISTING

<110> APPLERA CORPORATION et al.

<120> ISOLATED HUMAN TRANSPORTER PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
AND USES THEREOF

<130> CL001362-PCT

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1551
<212> DNA
<213> Homo Sapien

<400> 1
atggtgctct cccaggagga gccggactcc gcgcggggca cgagcgaggc gcagccgctc 60
ggccccgcgc ccacgggggc cgctccgccg cccggcccgg gaccctcgga cagccccgag 120
gcggctgtcg agaaggtgga ggtggagctg gcggggccgg cgaccgcgga gccccatgag 180
ccccccgaac cccccgaggg cggctggggc tggctggtga tgctggcggc catgtggtgc 240
aacgggtcgg tgttcggcat ccagaacgct tgcggggtgc tcttcgtgtc catgctggaa 300
accttcggct ccaaagacga tgacaagatg gtctttaaga cagcagcatg ggtaggttct 360
ctctccatgg ggatgatttt cttttgctgc ccaatagtca gcgtcttcac agacctattt 420
ggttgtcgga aaacagctgt cgtgggtgct gctgttggat ttgttgggct catgtccagt 480
tcttttgtaa gttccatcga gcctctgtac cttacctatg gaatcatatt tgcctgcggc 540
tgctcctttg cataccagcc ttcattggtc attttgggac actatttcaa gaagcgcctt 600
ggactggtga atggcattgt cactgctggc agcagtgtct tcacaatcct gctgcctttg 660
ctcttaaggg ttctgattga cagcgtgggc ctcttttaca cattgagggt gctctgcatc 720
ttcatgtttg ttctctttct ggctggcttt acttaccgac ctcttgctac cagtaccaaa 780
gataaagaga gtggaggtag cggatcctcc ctcttttcca ggaaaaagtt cagtcctcca 840
aaaaaaattt tcaattttgc catcttcaag gtgacagctt atgcagtgtg ggcagttgga 900
ataccacttg cacttttttgg atactttgtg ccttatgttc acttgatgaa acatgtaaat 960
gaaagatttc aagatgaaaa aaataaagag gttgttctca tgtgcattgg cgtcacttca 1020
ggagttggac gactgctctt tggccggatt gcagattatg tgcctggtgt gaagaaggtt 1080
tatctacagg tactctcctt tttcttcatt ggtctgatgt ccatgatgat tcctctgtgt 1140
agcatctttg gggccctcat tgctgtgtgc ctcatcatgg gtctcttcga tggatgcttc 1200
atttccatta tggctcccat agcctttgag ttagttggtg cccaggatgt ctcccaagca 1260
attggatttc tgctcggatt catgtctata cccatgactg ttggcccacc cattgcaggg 1320
ttacttcgtg acaaactggg ctcctatgat gtggcattct acctcgctgg agtccctccc 1380
cttattggag gtgctgtgct ttgttttatc ccgtggatcc atagtaagaa gcaaagagag 1440
atcagtaaaa ccactggaaa agaaaagatg gagaaaatgt tggaaaacca gaactctctg 1500
ctgtcaagtt catctggaat gttcaagaaa gaatctgact ctattattta a 1551

<210> 2
<211> 516
<212> PRT
<213> Homo Sapien

<400> 2
Met Val Leu Ser Gln Glu Glu Pro Asp Ser Ala Arg Gly Thr Ser Glu
1 5 10 15
Ala Gln Pro Leu Gly Pro Ala Pro Thr Gly Ala Ala Pro Pro Pro Gly
20 25 30
Pro Gly Pro Ser Asp Ser Pro Glu Ala Ala Val Glu Lys Val Glu Val
35 40 45

Glu Leu Ala Gly Pro Ala Thr Ala Glu Pro His Glu Pro Pro Glu Pro
50 55 60
Pro Glu Gly Gly Trp Gly Trp Leu Val Met Leu Ala Ala Met Trp Cys
65 70 75 80
Asn Gly Ser Val Phe Gly Ile Gln Asn Ala Cys Gly Val Leu Phe Val
85 90 95
Ser Met Leu Glu Thr Phe Gly Ser Lys Asp Asp Asp Lys Met Val Phe
100 105 110
Lys Thr Ala Ala Trp Val Gly Ser Leu Ser Met Gly Met Ile Phe Phe
115 120 125
Cys Cys Pro Ile Val Ser Val Phe Thr Asp Leu Phe Gly Cys Arg Lys
130 135 140
Thr Ala Val Val Gly Ala Ala Val Gly Phe Val Gly Leu Met Ser Ser
145 150 155 160
Ser Phe Val Ser Ser Ile Glu Pro Leu Tyr Leu Thr Tyr Gly Ile Ile
165 170 175
Phe Ala Cys Gly Cys Ser Phe Ala Tyr Gln Pro Ser Leu Val Ile Leu
180 185 190
Gly His Tyr Phe Lys Lys Arg Leu Gly Leu Val Asn Gly Ile Val Thr
195 200 205
Ala Gly Ser Ser Val Phe Thr Ile Leu Leu Pro Leu Leu Leu Arg Val
210 215 220
Leu Ile Asp Ser Val Gly Leu Phe Tyr Thr Leu Arg Val Leu Cys Ile
225 230 235 240
Phe Met Phe Val Leu Phe Leu Ala Gly Phe Thr Tyr Arg Pro Leu Ala
245 250 255
Thr Ser Thr Lys Asp Lys Glu Ser Gly Gly Ser Gly Ser Ser Leu Phe
260 265 270
Ser Arg Lys Lys Phe Ser Pro Pro Lys Lys Ile Phe Asn Phe Ala Ile
275 280 285
Phe Lys Val Thr Ala Tyr Ala Val Trp Ala Val Gly Ile Pro Leu Ala
290 295 300
Leu Phe Gly Tyr Phe Val Pro Tyr Val His Leu Met Lys His Val Asn
305 310 315 320
Glu Arg Phe Gln Asp Glu Lys Asn Lys Glu Val Val Leu Met Cys Ile
325 330 335
Gly Val Thr Ser Gly Val Gly Arg Leu Leu Phe Gly Arg Ile Ala Asp
340 345 350
Tyr Val Pro Gly Val Lys Lys Val Tyr Leu Gln Val Leu Ser Phe Phe
355 360 365
Phe Ile Gly Leu Met Ser Met Met Ile Pro Leu Cys Ser Ile Phe Gly
370 375 380
Ala Leu Ile Ala Val Cys Leu Ile Met Gly Leu Phe Asp Gly Cys Phe
385 390 395 400
Ile Ser Ile Met Ala Pro Ile Ala Phe Glu Leu Val Gly Ala Gln Asp
405 410 415
Val Ser Gln Ala Ile Gly Phe Leu Leu Gly Phe Met Ser Ile Pro Met
420 425 430
Thr Val Gly Pro Pro Ile Ala Gly Leu Leu Arg Asp Lys Leu Gly Ser
435 440 445
Tyr Asp Val Ala Phe Tyr Leu Ala Gly Val Pro Pro Leu Ile Gly Gly
450 455 460
Ala Val Leu Cys Phe Ile Pro Trp Ile His Ser Lys Lys Gln Arg Glu
465 470 475 480
Ile Ser Lys Thr Thr Gly Lys Glu Lys Met Glu Lys Met Leu Glu Asn
485 490 495
Gln Asn Ser Leu Leu Ser Ser Ser Ser Gly Met Phe Lys Lys Glu Ser
500 505 510
Asp Ser Ile Ile
515

<210> 3
<211> 139573
<212> DNA
<213> Homo Sapien

<400> 3
aatgggtatt ttgtagactg tccttgattg ggatttgtct aatgtttttc tgatggttag 60
ccagggatta tgggttttgg caggaagacc acagaggtaa agtaccattt tcatcacatc 120
atatcgggga tacattatca tctagttgag gtactgtgtg ccatttttg caccctaaag 180
ttatttcttc cccccactcc ccctttccat cctatactct ttggaagaaa gttactacgc 240
atacccacac ttaaagagta aaccattgta cttcacctcc atgagggagg gagtatgttc 300
ataaagtatt tacatttcct gcaggagaga tttgtctatt ctctcctcat tatttattta 360
atcatttact tacatcagta ctgactcgtg gataattctt acatatgtgt ttgtttgtgt 420
gcatgcaaat atataatcga tgtgctttct ttgcccaata atatgttgtg gacaactttc 480
aaagtcaata aatacagatg accttcagaa cttttagagg ttttaaagta agtatctaat 540
cagtcttcta ccaatgtaca ttatacttcc aaattttcct tatttccaac aatactgggg 600
tatcatcttc atacatacat ttttgtgcac ttatgtgcct attcctttgt ttactatttt 660
accctcattt ctaaggcaga ttacacttga gctatgttcc ccattccaca accaagcacg 720
gcttgctttc cttagtttat gctatttttct ctacctggaa tgctcttttc ctgtcttgac 780
ccactgaagt cgtatgtatg aatcagggct tcggccaaag gctgtttgct gtagaggctc 840
ttctacagtg tttggagaga atttaaggga ctatctcatc tctctctttt gtacatgtat 900
tataacacat catctgagcc tcctagtctc tcccaggact ctttttttcct accagtttat 960
caactgataa gaggcagaaa cgagatcaat cgcactcatc tgtgtactct atcacagtgg 1020
tgggcacatc aagtaactag catattttga ctttgattga agtgaagaat acgaataaca 1080
gaaattaaga agcatcctca atattgcata gcaggttact cttcttttct tttacatagg 1140
atggcactcc atgcttcagg gagacagagg agttgaatac aggttttagt ttttgtttaa 1200
agtgaaaacg actctgatgt agttgaaaag taatgctttc tagctgtctg ttaaaaatgt 1260
ttgttggttg aagacttcgg aattgcagtc cagtgaggac tgaaaataag catctttggt 1320
gtgccaaata ttcataagga aattgtatac gaatgcaaga gaatggaact gaagtaataa 1380
aataagggct ctgatccttc agatgactta tttaagaagc caggtggcat aacgaatctt 1440
acatattata attagtactg agaggtgaat gccaaaacat aaaacaaaca caatcgagac 1500
aatgttagtg tgactgtgac gctgtgtcgg tgagttgagg ctaacgatcc agtgtggctc 1560
tcctgaaggc ccaccgcgcc cgcacctagg agacgcgccc cttctgctca tgctttgagg 1620
cggggtgacc cacacatctg tgcccctctc tgagcaggag gaggccccgt cgcagacgcg 1680
cgcgcagaca gcgtctgccg cgggcacctg gggccgcgcg ccgcggggcg ccccgcctcc 1740
gctctccgag gcccaatcat ctggaggctg tgggggcacg tcccgctccc ggccacgccc 1800
ccagccggcg gggcgggggg tgcttttaag aaccggcggc tggcagtggg ctcagtcggg 1860
ggtgcggggc tgtgacctag aggcttcagt gtcgatcccc gaggtgttcg cgcgcgccag 1920
ctgtcctcgc ggccgcctgc gcgctggccg cctgcgcgct gccagcccgc ccgcccgcca 1980
ggggctccgc cgccctcgcc tcggcctcgt tagcccgcca ggagccccgc agctcctccg 2040
ggagcccgct ggtaactcgc gtccctcgcg cttctccggc gcctgagggg cccgcctcgg 2100
gccatggtgc tctcccagga ggagccggac tccgcgcggg gcacgagcga ggcgcagccg 2160
ctcggccccg cgcccacggg ggccgctccg ccgcccggcc cgggaccctc ggacagcccc 2220
gaggcggctg tcgagaaggt ggaggtggag ctggcggggc cggcgaccgc ggagccccat 2280
gagccccccg aacccccccga gggcggctgg ggctggctgg tgatgctggc ggccatgtgg 2340
tgcaacgggt cggtgttcgg catccagaac gcttgcgggg tgctcttcgt gtccatgctg 2400
gaaaccttcg gctccaaaga cgatgacaag atggtcttta agacaggtga ggcgcggcgc 2460
ccgccgaggc cagcctgggc gacccgcgtg gggcccccga gcgcatcccg cgtgtgggct 2520
gtgtctgcct ccgagtgtgc atgtcggtgg gtccctgtgc cagagggtgc gagcaggggg 2580
gtctttcgag ttgcagacag agcctgccgc ttctggggcc tcggggtgcc cgtctttata 2640
tggaatccag ctgcagagct gtgtgtttgc aagcaggtcg cagaacttac ttgccgagat 2700
cgtcctcctt tcccctcagc agagcagacg ctaacagtcc acaggagccc ttccttttat 2760
tgtttgaaaa caaacagaac cccagaacct tcaaccccag tcatcgccct gtcatttttg 2820
tggtctcttt cgtgactatg ccagttatgt agttcttcac ctgcctccct gggccgcaga 2880
ggggtgtgcg tatgttggcg gggcgggggg tggagtttgg aggaatgaaa gagatttgta 2940
cgaaggtcac tggagttcca aaggggggccc tgcaagagtc acggttccgt gcgttcccgt 3000
ccccccgcct ttttttgccc tctgggttaa atgtagaaaa cacgggaggc agccggatta 3060
gggactagga tgaggaaggt gaagggttgc ttcttccctc ttcctgttgt gtttttgac 3120
atttttttt aaccatatag taaattagat acaaaaggtg cagattcagc gttttctccc 3180

tgtagagcat tattatgact ttttggctgg ttaggcaaaa aacaaatcta agaccttctg 3240
catgacactt taacataaat tctttcactt tatcctgcaa ggtgagcgcg gtcaacccca 3300
tttgggtgag aaaactgtag ctcagtgaaa gtgtcttggt gggtagtaga atggcaataa 3360
aacacatatc aactgacttc aagggctaag tgatttccat tactaaatca acctccctcc 3420
ccatcatttg gggtaacttt atatgattaa tagtcttttt ttttaacctt gattttctat 3480
tatttttaga gtgaatattt cttaggtctt tagtatgcat atgaggaatg ggcaagactg 3540
taataaattc tgagacaaag gtaatgctgg gttatgctga gagttttaaa acctgacata 3600
aatactatta aactatttgt atcattctgc aacttacttt tcttccattc cgcatcatgt 3660
ttgtgactta tccacataat acctcagtgt gaactgataa ctcaaattct ttccatttta 3720
acttaggtgg tttgcattgt ttgactatat tatactctat gcattctccc tctgatgggc 3780
atttagattg cttccaaact cattctaaac aatgctgcaa tgaatattct tgtacactct 3840
cgttatgcat gtgaatacgg taccatttta acctggaatt tctgttcttt aaatagctat 3900
tgaaactgct gtggtatgcg ggtcaatggg ctaggtacaa aaagtgttaa aaatgtagta 3960
acatatcctt accatttaag ggaagtaatc attgtaaagt ttagcagggg agatatgcat 4020
atataatagc aaacaaaaat agtttgttgt cttttctata tgagtattgg gtgtcagaga 4080
gaaaagcccc aaaagaaggc agaattgaca gagttaacat ttaaagacta gttccaacat 4140
ttaccatatt cctgcctggg attacagatt ttttaatgca gtcaagataa cagcagtctt 4200
ttgtttatca ttgtttttgc aaattcagtt aagtagatcc tttggtgtct gtgggtgggt 4260
ttttttttt ttttttttt ttttttttt tttttttgag agagagagag agcaattgcc 4320
agagagacca tagctttgcc agggatgaga attttgcagt gtcaaagtct ctacctacta 4380
cttgtcccca aagttctaat tggctacaca atatcccaat actgggtagc tgagagtgag 4440
ggaaggaacc tggtttttct tttgcactct gtggaacttt gtgttttcca ttttgatgaa 4500
tatctttttt ctttttactc agttcagtct ttgacaactt tttcagtcat gtttgtgtat 4560
gtgtgggtat atatcatata aacagttgca caggtgtgct agttaaatgt gtgaagatct 4620
ttgtgtttct ctgcctgact gctgtatatc tatttatggt tgtgccattg cacaagggtg 4680
cccaactcag gggtaagtgg ggactgaaaa ccagcctggg ctctgggtgc cctttgtctg 4740
attcctacag aagggcccta tgcacttgta atggccctgt ataacacagc atctagattg 4800
aacaatggcc attacttggg tgctaggtaa tacatatatg actgatagat gttaagggct 4860
gaggaagaaa acagatttaa acttagtgct gaaaaaatgg ttacaagata gtctttaagc 4920
cagttattgt tgagatctct ctctcttccc ctgtccccta ccccttctc tttccttcag 4980
tgcacacaca cacacacaaa ggtgtttcat gaagtccctc atctaccaca gtcactgtta 5040
tttgagaata tctgctttga agtttgattg gtcacacttt ttcactttga tattcgaatg 5100
ctgagtcgtc tgtgatcaag catatgcaag cttcaaatac atgccaaaaa atatctggaa 5160
tttgtttaag ccttttattt ttcaaagttt tggtctattt tctattaccg tactcatgat 5220
ggataatcct ggtgttagag tacagctagt tctgtctcct tgtttccatt acttctttat 5280
agcaagtgac tagcctaagg atatacaggg aggtggtggt ggaatggaat ctaggtctcc 5340
aaatgatggt gcgcatttct tgagtacttt cctgtggcta agcactttag atgcgttcct 5400
atttaaacct taccacgatt ctctgataga ctttgttaat atctttcttt tcagatatgg 5460
gaactcaggc ttacagagtt taagtaagaa gtggagccag aattcaaccc caggcttatc 5520
tgactctaag agctgggatt tttattttaa ttatttattt atttaaaata tggaatgctt 5580
catgaatttg catgtcatcc ttgttcaggg gtcacgctaa tcttctctgt gtcattccaa 5640
ttttagtaga ttgttgttca aagtgctgct gaagcaagca ccaggagctg ggttttaatc 5700
attcatcata ttgcattgac tagataacat tctgcaaata cgatgttttt tatgttgttg 5760
attaatttaa gtgttagtga ttggttgagt gctctaccat gcattctggg attagaaaga 5820
agggtccctg tttcttggtc ctactttgtg gtgaataaac aattgcaaat tattaatgtc 5880
tcaaactata tttctgaagt gtagagagac ttccatagaa gaacaagata cttccatatg 5940
ccgttcaagc aaaagtctgg ggtttccttt gaagaacttt tagattgatc cacagcagga 6000
caatgtttct aggcagaact gaggaggagc ctttcttagg ctcacttctc ttcagggctc 6060
tgttaactct tcccacgcaa tggataatct acccaaaatt tctcaggaaa gggcctgaag 6120
aagttcattc acactaaggt gtaagtgagt ttacacatct tactgttaat tctctttata 6180
caaatgttta ccaagttatc taacacgctt tgttttgggc tctgtcctgg ggactggaga 6240
taatgactga gagagaaaat gtcagctgtt tcaaagtagc ttaggatctg ttgtgggata 6300
caaattaata acagaccaga agtaatagaa tatttccctg aaggattttc aatataacag 6360
gactcagttt tactataaaa ggctgaaatt ctaaggtcat ttcaacaggt ggtggggttg 6420
ggggtgggga aggcatttga cgcctctttc tctatggtta taaatctcac ttggtgaaat 6480
taagactttg gaaaggggaa gtaagccaac tccaagttgg gcagtagaac caatgaaaaa 6540
tgctgacggc atcacagtcc cattatggtg cccagctgcc aatgacatgg cactcagagg 6600
agtgtctcac acatactgct ctgtctgagg gagcaagcta agcttgagtt gtctcttttt 6660
ttgttgtttt ttttttttt ttttttttga gacagattct cactctgtcg cccaggctgg 6720
agtgcagtgg caccatctcg gctcactgca accactgcct cccggatgca agcaattctg 6780

cctcagcctt ccgagtagct ggactacctg cgcttgccac cacacctggc taatttttgt 6840
atttttagta gagacagggt ttcaccatat tggccaggct ggtctcaaac tcctgacctc 6900
gtgatccacc tgcctcggct tcccaaagtg ctgggattac aggcataagc caccgcgcct 6960
ggccaagttg tctcttttta gttgaatttt tacctgttca catgtgtatt cttcttgcct 7020
aggtagagag gaatcagaca ctctggggaa gaatacaaag aaatacaatt aagtggaaca 7080
ttgtttttct ttagaaagtg caattttggg ctgggcgcag tggctcatgc ctgtaatccc 7140
agcccttgg gaggccaagg caggtggatc acctgaggtc aggagtttga gaccagcctg 7200
gccaagatgg tgaaaccccg tttctactaa aaatacaaaa aattagctgg gcatggtggc 7260
ggatgcgtgt aatcccagct attcgggagg ctgaggcagg agaattgctt gaacctggga 7320
ggcagaggtt gtagtgagcc aagatggcgc cactgtactc cagcatgggc aacaagagtg 7380
aaactccgtc tcaaaaaaaa aaaaaagaaa agaaaaaaag aaaaaagaaa gagcaacttt 7440
gttttaactc tgctagatac tggaaaaccc atggaactaa tgaagagcct agggcttttt 7500
atttgttttg agattgtgcc atttcactcc agcctgggca acaagagaga aactttgtct 7560
cacacacaaa aaaagtgtaa atcaaaacat taaaaattaa gtagtttgga agtagattat 7620
caaaaaggtc ctgaaaggga ggttctttgg ctataatctt taacgcaact ctacactccc 7680
tgtatggaga cagatttctt tttagatggt tacagtcaca aagtagggtt ttcagtagca 7740
tttagggatg aatgaatctt gcagcacctc tccatgtatc ttgctagccc ctctgaaact 7800
tcaggtcagt tagtgcttcc tcagaaattg ttccccccac accaagtttt cacatttaca 7860
gttatactga tatccacatt gtactgttgt atgtgacacc tagattatag gaaattttgg 7920
ctatagttca gaaattaact gctatgtttt gcctttacgc taaagagatt ttgttttgtt 7980
tagtaggaaa agcggcctgc ataactagcc atttctgtat cttagaaaaa tttttagtaa 8040
cagtcctttg ttgagctagt tacagtgaac aaataatctg gttcatggtc ctatacatct 8100
ttcactataa gaaaaatacc tgattgttat ttacactgga agagaggtag aaaagctaag 8160
agaactcact tatggcaata aaccaatcta aactacctgc taaaataagt gagaagatta 8220
taaaaatggt tctaggattt tggaataata gtgagtatgg tatgggcgtt tcatacttca 8280
tttccagaat gtttctggat taagtgcgag actgaatagc atatatagtg aattctaatt 8340
aaatacaaca atgtgagatt cctgtggtgt tttttcatgg aattaaaaat taataatttc 8400
aataaaatta actgctgaaa gaacccaatt agccaaaatg aaaagcataa cacatttttt 8460
caggagcgat tttgaggtgt cttttagaat aaattgtact ctgcttttga tgtgatttgc 8520
tacatctttt tgttgcagtt ccttgaggct cagcccctgg ccatatactt gcttcacttt 8580
tcctgctttc ttccatccac tgtcttgggg ctgttatttc caaatctcat cactgtgttc 8640
aagacttatt tactatttct ggacagttcc atttgggtac acaggtacat catactaaac 8700
taacatgaac tcatttttca gctacaccat acaaccttcc cttactacca aaaatgacag 8760
cccatttgtc ggcattcttc tgaatccata ttcctccttc ttaattctct gtgcatgata 8820
cctctggttg tttaagtcag aaacctggaa tgtatcttag cttcttattt ctcttccttc 8880
cttacctgtt ttgttaaatc agagttcttt tgtgtttacc ttcttaatgc ctcataaatc 8940
agtcccactt ttctttcact actgtgtctt aattcatgcc ttcatttact ttttatttta 9000
ttttttggga tagggcctca ctgtgttccc caggctggag tgcagtggca tgatcatagc 9060
tcactgcagc ctcaaacttc taagctcaag caatcctcct atctcagctt tttgagtagc 9120
tgcaacacca ggcacatgcc accatggccg gctaattaaa aaaaattttt atgtggagat 9180
ggcatcttgc tattttgctc agccttgtct tgaactcctg gcctcaagcg accctcctgc 9240
ctcaacttcc caaagtgttg ggactacagg cgggacctac tgtgcctggc caccttcatt 9300
actattggca acaattagtc ataacccctt aacaggattg cttgtcctca gttgtacatc 9360
tgagtgattt ttctaaaaga ttggaccata tgatttttct gtttaaatgc ccagtgacac 9420
tcattacttt taggaaaatg tcaaactccc tactccgaag gcctgcaagc tctggccctt 9480
gcctggccct ctagccttgc ccctgcttct ctcccttact ggtctttgtg ttctagccaa 9540
cctgtaggtg ttacactggc ccaaatttgt cttgctgctt tttgcctctg taccttttgtg 9600
tgtgccactc ctgtcttcag tgcgatggtt ggtccttgtg agattctgat gaaatggttg 9660
gccatttttat tcttatgtca caatcctggg acacaaacag tgattttatg caattgttat 9720
gtatttgatg cacttggaat attgggggta gctacatttt ggagttttga gaacgaattc 9780
aaataagtta caaattatgt ttaaagtggt agacagagaa cctgatttca acctattcta 9840
ataaagcatt ccgtgaaagc cattttaaag atgatccata tttgttaaag tggtaatttt 9900
tatattctct gatatggttt ggctgtgttc ccacccaaat ctcatcttga attgtagctc 9960
ccataatcca cacttgtcat gggagggagg taattacctg gtgggaggta attgaatcat 10020
gggggcaggt tttcctgtgc tgttcttgtg atagttaata agtctcatga gatctgatgg 10080
tttataaagg gcagttcccc tgcacacttt ctgttgcctg ttgccatgta agacatgcct 10140
ttgctcctct ttcaccttcc accatgattg tgaggcctcc ctagccatgt ggaactgtga 10200
gtccattaaa cctctttttc tttataaatt acccagcctc agatatttct tcataacagt 10260
atgaaaatgg agtaatacat tccattacca taaagaaaag gctttcatgt acattatttt 10320
ttagagtagc cttgtggtat gtcattacct ccatggatag ataagaaagt tgcaacttgc 10380

acagtattag gattgatatc agtatttact tttattaagt tgaacttaag agcagctttt 10440
tggctggaaa aaagttgtac ttatgtcaaa gttgtcctga aagtagaatc ctactcctgt 10500
ccccagcctg aaactattta ctacatattt acttgcatgt tctttagaat attctctcaa 10560
tagtgtctcc tactcaagtc atcagaaatg ctgtgatgtc atttttggaa aagaattcca 10620
gagttatcac cgcctagcta gaaatctggt cttatattca aattaaacaa gcaaacctta 10680
acaaaaacaa gctaaaccct aaacataaca tgaacagtca gctcaccaca gttctgagca 10740
cctgccttgg cctggtgcca cccacgcagg gactgtggat gtttttattg gcagagattc 10800
agaacaggaa actccagcac acctgggaac tgcgcagacc caccacataa gacagatagc 10860
ctatcagtgg ctggaggaat ggaggaaagc agtgctttca aatgtacatg ccaaatgtgt 10920
atgatcatac ctctttgtta aagtgccttc tttaacagca aaagtaattc ctcaccttgc 10980
atataggaac taaaaaaaag tcgatgaaga aatggcttgc cttattttca agtaagaagt 11040
ctttttttcat ttcactaatt tttaattatg ggcataagta tgaaatacag attagaaata 11100
ctgaatgtgg accaaagcaa tgtttccttt gtggaccaaa gcagtgaatc ttttttcttt 11160
ctttctttct tttctttttt ttgttaagag acagggtctt gctctcttgc tcagactgga 11220
gtgcagtcag tgatgtgatg gctcaccata acctcaacct cttgggctca agggatcctc 11280
ctgccccaac ctcctgaaca gctgggatta caggcacata ccaccacacc tggctaattt 11340
ttaaaaattt ttttgtgggg gagggtctct ctatattgcc caagctggtt tcaaatgcct 11400
gagctcaagt gatcctccaa cctcaggctc ccagagtgtt gggattacag gtttgagcca 11460
ctgtgcctgg ccccaaaaca gtgtttgtaa ctccctttgt cccctcttga taaacataaa 11520
agtctcatgg tacttcagaa ttttgttgca cacatccaag tgtagtttgc ctttccttgt 11580
gagtggcaga agacaatgtc atactctgta tttatccatc agccaaaatt ttgtcaagct 11640
ttacttttat tttttaaatt tttatttgta tttttttttg gagacagagt cttgctctgt 11700
cgcccaggct ggagtgcagt ggcgtgatct cggctcactg caacctccga tgcctgggtt 11760
caagcgattt tcgtgtccca gcctccagag cagctgggac tacaggcacg caccaccatg 11820
cctggctaat ttttttgtat ttttagtaga ggtagggttt cgccatgttg gtcaggctgg 11880
tctcgaactc ctaatctcag gtgacccacc cgcctcagtc tcccaaagtg ctgggattac 11940
aggcatgagc caccatgccc ggccaagctt tacttttagt tatgttgtgg atataattag 12000
aattattttc ttgttcttta aatacagaat tataatattg aatgtgtgct ttaaaaaaat 12060
tagtaaatgt acccctctag aacttctgat ctatgcctct tagtgagtga ggggagctgt 12120
gccgctttcc tcctccagtg cctccactta agaatcactc actaaggagt tttaaattca 12180
attaaaggta tcctttagat agttaagtct aaatgaaagg tcaatgatta aattaatgga 12240
taaaagtcca ttgcacctac ggaaaggtgc attggtctac agttcagcta gtacctattt 12300
tttgctaaag attaaattgc agctgggcgg ggtggctcat acctgtaatc ccagcacttt 12360
gaaaacctga tggaggccgg gtgccgtggc tcatgcctgt aaatcttagc actttgggag 12420
gccgaggtgg gcggatcact tgaggtccgg agttagagac cagcctgacc aacatggaga 12480
aaccctgtct ctactaaaaa tactaaatta gccgggcatg gtggcgcatg cctgtaatct 12540
cagctactgg ggaagctgag gtaggagaat cgcttgaacc cgcgagttgg aggttgcaat 12600
gagccaagat catgccattg tactccagcc tgggcaacaa gagtgaaact cagtctcaaa 12660
aaaaaagaaa accctaggtg gaaggatcgc ttgagtccag gagttcaaga acagcctggg 12720
caacatagtg agaccccatc tctactttta attaaaataa aattaacatt aaaaagtgta 12780
gtaaaatttt taaaaagagt aaattgtatc agcagtgttc tgcctgtgtt cagaaagcca 12840
aaatttatat ttgtgtttca tttagattgg atccaagacc aattttgaga tgtgttttaa 12900
tattacaaaa atagaaaact acctgttcat taaatggtga tatttattga cttttctgtg 12960
tagataagta ttaatgccaa gtcaagtagg ttagtctgga gactgttttt attaataaaa 13020
acctttttcat cttaattatc ctttttatta gttttgactt tatgttgcaa cttcaaagca 13080
gcatctcaaa gggttcacat ttgtcagtgg ttgtttaaac aacgacattc cacaaagtta 13140
attgatgttt tagtgtgaat ggggcaggaa gttgtcattg ttgtcctgta gtaacgtagc 13200
actcaaactt tgacaggagg accccagcca tgtagttatt gactgtaagt acacttccct 13260
taagagtaag ttgtttttta ggaataaatg ataaatgtat ggagttttag ttgtaggtgt 13320
tatgttttgt ctccttcttc tcagagaaaa atgttactgt ggaataggct taaactgaaa 13380
aaaggatctg atttttaata agatcagatt ctttggccat attttgatat tggttcaaaa 13440
caaatgttta atatcagatg cacaatgtta agagctctat aatgtagtgg taacacctga 13500
gcctcagcca gtgactacac attaagttct tgttcatttt ttgcaggggc aggggagctg 13560
gtgaggcata aagaagggtc agaggagata atagacttat atttgtattt ttatgcataa 13620
ttatatagga aacattaaat ccagagttga taaataattt gtatgtatgt atttatttat 13680
ttttgagact gggtcctgct ctgctgccca ggctgaagcg taggggtatg aacacagctc 13740
actgcagcct tgacctgggc tcaagcgatt ttcttgcctc agcctcccga gtagctggga 13800
ccacaggcat gtgccactac acctggctaa tttttaaagt ttttttgtag agatggggtc 13860
tcaccacgtt gcccaggttg gtcttgaact cctgggctca agcaatcctc ctgccttggc 13920
ctcccagagt gttgggatta tagatgtgag ccaccatgtt cagctgataa ataatttcta 13980

atctaaaaat cctattttgt atggagaggg gagggcaaat aggctatttt ttccacattt 14040
tgttgctggc cagaatctca gagggttttt acctgcatta aaaatgatta aggctgggca 14100
gaatggctca tgcctgtaat cccatcactt tgggaggctg aggcagatga attgcttgag 14160
cccagaagtt ctagagcagc ctgggcaaca tggtgaaacc ccatctctac gaaaaatgca 14220
aaaattagct ggacgtggtg gcaggcacct gtagtcccag ctactcaaaa ggctgaggtg 14280
ggaggatcac ctagctctag aggtcaaggg tgcaatgagc caagattgct ccactgcccg 14340
ccagcctggg cagcagagca agaccctgtc tcaaaaaaaa aaaaaaaaaa ggttattttt 14400
tttcagcata gacaaagcag gggaaggaaa agattatgtt tcaaatgttc atttaaacta 14460
ctacatttaa agtaatactt cctaatgatt taaactttag attagtctat ttatgggtca 14520
cctggaagat tctttataaa acatgagagt ttattacttc tttcaataca cgggtgtctg 14580
taaatgatgc tcaatagatc tgaagcctga actttctgaa gaaatgttgt gaaattatct 14640
atggatgtct acttgaattt aaataaaaat aaattgtaat acatttggtt ttattggttt 14700
tgaatttgta attatttggg gatttggatt tggtttatct acagttgtct ttttttttga 14760
ggtggagtct cgctttgtca ccaggctgga gtgcagtggt gcgatctcgg ctcactgcaa 14820
tctctgcctc ccgtgtacaa gcgattctcc tgcttcagct tcctgagtag ctaggattac 14880
aggcatgcac ctccatcccc agctaatttt tgtatttta gtagagatgg ggtttcacca 14940
tgttagccag gatgatctcg atatcttgac ttcgtgatcc gcctcggcct accaaagtgc 15000
tgtgattaca ggtgtgagcc accgtgcccg gcctacagtt gtctttttt actcactccc 15060
acagatgaat cattataagg aggttagctt tccttaaaga atacactctc cttaagcggt 15120
tttatcaaca aagccaggga atgccaaact ttaacacttt tacctaaatt tataactgat 15180
gcatgtatgc atatatacat acatacatac atgtatatgt tgttatatat gtatgggtat 15240
catcagtata gtctatcagt atagtaattg tttatctgaa acttggggtt ctctctcgct 15300
ctttctctcc ctctctctct ctctctctct ctctctctct ctctatatat atatatatat 15360
atatatataa atatatgtat atatattttt tctttttttg agacaggatc tcattctgtc 15420
acccaggctg gagtgcagtg gtgggatcat ggctcactgc agcctcgacc tcctgggctc 15480
aagtgatcct cccacctcaa cctcccaagt agttaggact acaggggcat gccgctacac 15540
gtaactaatt tttgtatttt tttgtagaga cagggttttg ccttgttgcc caggctggtc 15600
ctgaaatcct tggctcaagc aatctgtcca cctcagcctc tgaaagtgct ggcattacag 15660
gtctgagcca ctgcgcccag cctagatttt tttgaattgt aaaaaagtaa cctgctccct 15720
actgaagtaa atagagttaa aaaaagtaat ctggtacaga cacctgtatt ttctgacacc 15780
cctagaagag tcccaggtac cctataatca aatacattaa catttctgca gcaaaatgta 15840
tggataagtg agttaaatag agaccatgag tagcttcagg tcagttcaga tcaagttttg 15900
cttctaatta aatgttgata ttctcttaca aaaactttgg gtttgggttt tcagatttgc 15960
aaataaataa ttataaatta ttatttttt tgagacagag tcttgctgtg ttgctcaggc 16020
tggagtgcca tggcacgatc acggctcact gcaacctcaa cctcaggctg aagccatcct 16080
cccacttcag cctcccaagt agctgggact acaggagtgt gccaacatgt ccagctcatt 16140
tttgtattct tagtagagac agggtttcgc cgtgttggcc aggctggtct caaactcctg 16200
gtctcaagtg atccgcctgc cttggccttc caaagtgttg agattatagg tgtcagccac 16260
tgggcctggc agaattatac atttatatgt caatatttgc ttttgttttc tgtttttcag 16320
taaacgtttt ttaaggtaca tttttctgta tctcataagg cacctgctta attgtttcag 16380
taagtgtgat gttctaccat attggtctac cctaggttac tcaaccaggc ctcctttgtt 16440
tagtgagtag caggcagtgt tgtacaacat atgtagcata tctgtatatg tcgtcgaaca 16500
aattgttttt ttccccctctc ttggattgct tccttgggtg tacgtccaga agtgagatta 16560
ctggttcaaa gggtatgaac aactttataa cacctgttac acattgccag attattcttt 16620
agaaaacttg aatcagttta ttgtgccacc agtgatgtgc tggcttcctg aaaaccctac 16680
caatgtttgg ttttatttttt attagtattt gctaatttga taagtactaa tgatattttt 16740
taaaagtagt ttaaaatcat atttcagtgc ttataagtct gtgttcccag tttttgagcc 16800
ctttagaagc tgcaaatgac ctggcaatta tatataatat ttgaaaatac aagaggacat 16860
atgccagtga atatattaga gtaaaacttc attcccatag gtaatgaagg aatgcttgag 16920
attatcttag gccttagatt ctcacctgac acatcttggc aggtagacca tgtccttgtt 16980
tcctctgctg tcttagccca ggtgttgatc aaggtctgtc ttagggcggg ggataggaat 17040
ggaaataaac catgtagaga cttgggcatg aggactttgt gattcttcca ggtgacatct 17100
catccttcag aggatcaagt ctgcaagagt agccatatct taatctcttt cagtgctatc 17160
accttgcatc aacctctgga ctcgagctaa ttccgttgaa aatattttat taattaattt 17220
tggggtatgt taaaaatttt gtttgcatgt atttatttag ttatttttat gagacagggt 17280
ctcgctctgt cacccatgct ggagtacata cggttgcacg ctcatggctc actgcagcct 17340
tgacttccca ggctcaagtg atcctcccac ctcagcctcc tgagtagctg ggactacaag 17400
tgcatgtcac cacatttggc taattttcat atttttttgta gagacggggt ttcgccacat 17460
tgcccaggct ggtctcaaac tcctggactc aagggatctg cctgcctcag cctcccgaag 17520
tgctgagatt ataggtgtga accaccgggc ccggctcccc atttaatttt gttgttgttg 17580

```
tttttagat ggagtttcac tcttgtcgcc caagctggaa tgcaatggca cgatctcggc   17640
ttacggccac ttccacctcc tgggttcaag cgattctcct gtctcagcct cccgagtagc   17700
tgggattaca gacacacgcc actatacctg gctaattttt gtatttttag tagaaatggg   17760
gtttcaccat gttggtcagg tgggtctcaa actattgacc tcaggtgatc cacctgcctc   17820
ggcctcccaa agtgctggga ttatagatgt gagccaccat gtccagccac ccatttaatt   17880
ttttgagcac aaaatatgta ctgagagcca cgcaagaaac aaattcgact tattccatgc   17940
tcttgagagg tcatgagggg aaacaaaatg atacataagt aactctgaga gaatatgctg   18000
cacatgctaa atcctgtgca agtaagatat aggattttag aggaagggag aatgacttct   18060
gattgagctg actagagaag gcttcagttt ttgagttagg tgttacgaga ttgggagact   18120
tttctcagca tatctaacag aagagggtat ccgaggtgag agtgtaaggc ctgggcaagg   18180
gttgggaggc agttctaata ctgaatgttc tgactgtggt ttactatgta tttcaggtta   18240
ttttgtttaa tctatccagt aatcctttca tgtaacaatt atgatgtgtg tgttttaggt   18300
ggggctacta aggctagtaa gtagtgaggc tggatttaaa cttaagtctc cagcttcgtg   18360
gcccaggttc tttatacttg actccacact gggcttatta agtgaatgac aaggagtttg   18420
atttgtcgag ggccgaggat atgtagtgaa aggtcataga ctataaggct ggcaaagaat   18480
ggtggagctg gggaacggaa gatctaaatg ccaaggtaag cttggacatt atttaataga   18540
ggcaatgggg agatactgag gatttctgaa gagggtttct aagtagtgct ttaaaaagat   18600
ctctgggttg cagaggcaag aaataggagg aacaccaatt acaaagcccc gaccgaaagc   18660
cgagtgagat ttgagggtca gccatggaaa tgggaatgaa gggacagatg tgagacattt   18720
caaaggatag actgcaggaa ttctggcaaa gaatgggttg gacagtgatg aagaacacag   18780
aaacagcatt gatgcctaca tctggagcct gtgtcttaat aatttttgat cagggagacc   18840
tacaatacat tttgtattgt gccccacaaa atcatgtcga cagccctaga tgatgtcatg   18900
gacaggtatg gggcatgagg agaaagagtt gccagggaaa ggctggctca cattttagac   18960
atgtggagcg ttaagggtca gcaagacatc aagggcgatg tcaggcaggt aactggaaat   19020
gcaaactggt agtctcagag tgacgctgag tccagtggtg gggctttgga aagaacatgt   19080
atctctaaga ggaacatgac ataatggaaa cccagaaatc ctgtcttgag cggactcagg   19140
ggctgcttct aggtaattta gttcatttct actgaaatca ttatatttaa aagtatggcc   19200
gcactgagat ggccactgta gctgctgcca cctcttagct ttggtcttaa aaaaaaaaaa   19260
aagtaataga acttccttaa aatgtctttc tagcctttgg atttctctaa ttcagatttg   19320
cgtcttccca agggtcaaaa ttatattttt actatccctg tcttaggtat ttccaaaact   19380
tcgtcttaag acttagtcat ttttttcctt caatttgaca tgactgctaa agacttttgg   19440
catgttcctc ctcctttcat ttgtgatgta attaagttgg tctgtaagtc ttatttttaa   19500
gatgttctag accaagagac tgtgagaata gcttacagtc atttcaacta atttatgtat   19560
tttaaattta aagtattgac agtggtgaaa acctgttcaa caagcagatg atgttatctt   19620
atatattcac agagtttagt aactgagcca actacttcat tcacagttca aaatgaaaac   19680
agctaattct tttaagtaag tatagattct actctttaaa agagttatct aggagagcta   19740
ctataatact attatagaat aagtgaaatt aagttatctg actgggactg gggtggaata   19800
ttgcacagtt tattgagaaa tatggactgt ttttgcacat tcataatgga catttgaggt   19860
ttgttagggg aggaggtgtc atctttatgg cactttctgg ctgggaaggg agtcagtcct   19920
aattgagata ataactagcc acctggccac acacaagtgt gttttgcctt agttacctgt   19980
cacacactga gcagtgagac tcaagagagt gtcaaagtac ttttcaatgc ataaagcact   20040
acagatctgt ccacactgtt gtgagtgagc aggttggcac ggtgcctgtg tgcgggcgtg   20100
tgtgttgact cacgtgctgt cctgtgatct ctcaggactc aggtcctgaa ttgctctgtt   20160
gtgactgaag cccagctgaa ggtgctggaa gtgcagtgac cctggaggaa gaaccagtaa   20220
cagcagaggg tggatgaaag ggaattgata gttgtgtaag aatagatatc tgccgttttt   20280
tgtaagccaa gacaccttta cccttccagt aattgtttca tcttttaata tcatttggct   20340
tcatttacag aatctgtatt aaagcaaaag tcagcatgta aggtggtatt ttgaccacat   20400
ttgtctgctg ttgtgccttc tgggtgaact gaagtactgg cttactgact agtaaatatg   20460
ttttctgaca attataggga agggaagaaa aggaaagtcc aattaaagca tttttctcct   20520
cagagtttga aaaatagaat tcatgcaatc ttttaaattc catgccaaca catcagacaa   20580
gaagagactt gatagtagta aaggttggga atcaaaagaa caatgtaagt ttttgatatt   20640
gacttcaaaa catggtgttc tataatttag tgttcatttg ttacgtgtat ggtattataa   20700
ttaattttgt atatgtggta gttatttttt gtactttgat taggaaacaa atattgagcc   20760
acttcaagag gcagactatt ttggaaaaaa aagtctggta aaagtaatgc ttaatctaat   20820
taatctgtct tcatcctctt aattcatcaa gatgactttg ggtgtctggt ggcaactgag   20880
aatgggttat ggaagaacgt caaggcaatg taatccctat tatttacggt tacttgagag   20940
gataaattaa tcagcggtca ctaatctttg gataatcact ctattgagct ggaactatcc   21000
ttagtatttt ggaaagcaag tcagtgagtt agaactgtca aaactgatca gcttttctaa   21060
gcttaatgat aagtgaatag aaactagttg ccttcaaccc tttcctccct gcattgcagc   21120
atgatcattc tgtaactctg gaaatggttt atggaacaac agtgaaaata cattgataca   21180
```

ctgtcttgtg gtagattttc agataggctt tagacaaagt tcagagcctt tcctctagct 21240
ggggattaac aaagctgcct tcatagttaa atgtttgcac cctgtgtatg cattttcagt 21300
tactagaatt aggtaagtta gtgttataaa attggtttga gtgtggattg tttaggaagt 21360
gagtcttttg gtggcagcaa ttctgttatg cattaaatag atacatattt tgaactagtc 21420
gacattgttt cagtctgtat ttattagatg ctggggtggg tatgggaata aagaaacgta 21480
tgaggggtct tggaaaagtt catggaaaaa atgtacacta tgaaaaaaaa ctgtgcatgg 21540
atttcaaaat atttttgaac caaaatcaac ttgtactaac ttgttacaac atgtctgaac 21600
ttgttacaac atgtctgaac agttcgagac attaagaaat gatatcgcac cagtttttaa 21660
aaagcgccta tcagggtaac atgaattctg ctaaaattga agcaagaaca aacatcaaat 21720
ttatggtggt gcttgggtag aaggatggtg aaatcattga tgctttacaa aaagtttatg 21780
gggacaatac tctaaaggaa ccagcagttt acaaatggct aacatccttt aagaagggac 21840
gagatgatgt tgaagaggaa gcccacagca gtagaccatc cgtgtcaatt ttcaaggaaa 21900
aaattaatct tgttcatgct gtaattgaag agggaacttt acatgaaatt ttaaacaagt 21960
gggattgaga tcctgtggca tatgtctgaa ggattgtaat aggagaggaa acatggcttt 22020
accagtatga tgctgaagac aaagcacaac caaagcaatg gctaccaaga ggtggaagtg 22080
atctagttaa agcaaaagca gactagtcaa gagcaaaggt catgataaga gacttttggg 22140
atgctcaagg tattttgctt gttgactttc tggggagcca aagaatgata atatttgctt 22200
attgtgtgtg ttttgagaaa attagccaaa gttttagcaa aacaaacaaa tacccaggga 22260
agctttacca gagagtcctt ctccaccagg acaatgttcc cggtcatcct ctcatcaaaa 22320
aagggcaatt ttgcaagagt tttgatggga aattattagg cattcactta cagtcttttt 22380
ttttttttta gttggagtct tactttgtct cccaggctgg agtgcagtgg tgcaatcttg 22440
gctcactgca acctccacct gccaggttca agcaattctc ctgcctcagc ctcccaagta 22500
gctgggatta caggcgtgat ccaccatgtc cagctatttt ttgtattttt agtagagatg 22560
gggtttcacc atgttgacca ggctggtctt gaattcctga cctcaggtga ttcacctgtc 22620
tcggccttcc aaagtgctgg gattacaggc atgagccact gtgcctggcc tatcttacag 22680
tcttgatttg gctttatctg acttcttttt ctttcctaat cttaaaaaat atttaaaggg 22740
cacctatttt tcttcagtta ataatgtaaa aaggactgca ttgacatgat taaattcctg 22800
ggaccctcaa ttctttagag atggactaat ggctggtatc aactcacaaa agtatcttga 22860
acttgatgga gcttatgttg agaaatgaag tgtatatttt cattatcttt taatttcatt 22920
ctttagtgaa ttttttgagg tccccttgta tacattttaa tcctaaggga ataaagaaag 22980
gaggaagtcc tagccctgtg ctgtctgcct aggtacagtg tctgaaacac agaccagtat 23040
tcacccttg aaatttgagg tttccattca ggaggttctc aaagagaata aatgagattg 23100
ctatgcaggt ggaatcaaag agcacacggc ttatttatca taatcaaaat aatgccattt 23160
tcataacaaa cttcacctgc ttatgtacat tgtaaattgt tgccttgata agcttcccgg 23220
agataaagta attcagctaa gtattatttc caatcataat tttgtgtcat tatgagcaac 23280
acaatactat atatgggatt gattcactgc agaactggaa taaatataaa ttagatcttt 23340
agaaaagaaa cgtagattta aaaatcttat gttagaaggc tcaattaatt aaatgtaatt 23400
aatttttaa aatcagcttt attgagggat gacttagata ttatataatt cacaaatttt 23460
aagtgtacag tttgatagtt ctgacaatca aactgtatac aatcatgtaa ccaccatcac 23520
aatcataata tagtgtgtcc atcaccccag ggtgtaccct cgtgatcctt tttgcagtta 23580
gtctttttcc cttacattct ggctcctgaa aacttgatct gctttctgtc actatagctg 23640
tgccttttct aaaattttat atgaatggaa tcatacagtg tgttttcttt tgtatctgtt 23700
tttcactcag catgatgctt ttgagatttc tccttgttgt ggtatgtatt agtagttctt 23760
tctttttat tactaagtag tattccattg tatgcctatg ccacattttt ttttttttt 23820
ttcgagacag agttttgctc tgacatccag gctggagtgc agtggtgtga tcatggctca 23880
ctgcagcctt gacttcccag actgaggtga tccacctgcc tcagccacct gagtaactgg 23940
gaccacaggt gtgtgctagt ctgtctaatt tttaaattgt ttgtagagat gggggtctct 24000
gtatattgcc caggctggtc tcaaactcct ggcctcaagc aatccttctg ccttggcccc 24060
tcaaagtgtt ggggttacaa gtgtgagcca tcacacctgg cctaccacaa ttttttatcg 24120
attcacatat tgatggatat tgggttgttt tcagttgttg cctattatga atagaactgc 24180
tatgaacatt tgtatgcaaa cctttgttgg gatgtatgtt tttatttctc ttttgtacat 24240
taaatttaaa tttaaatttt gttctgtatt atttgtattt ttaaatttct caagtgggta 24300
atactgtgta ctttttttt gaaattaaaa aaattgtggc tgaacaagga gataaaaaag 24360
taggagtgag aggactctgg agagttacag ggctttggtt tagaggattg gatgaatagt 24420
ggtgctgcca ataaagaaat ttaaatatgg ctgatatttc ctatatttaa gaaagaccaa 24480
agagggtcca ttgaaatgag tcagtgggaa atctctgatg acttcagcca gcacgctttc 24540
atcggcctgg atatatggga agtgaagtct gattatagtc tgtggagcag tgaatgggag 24600
gaagagatag gggtacaggc taagaaggga ggaagtcaag tcaaagggag aagtagggtg 24660
gtagctagag gaagattaga gtcaagcgag ggtaacaatt tttttttttt ttgaagatag 24720
gagtagcttg agaactaact taaagaagga gcctgtagag agggaggagg tgaagttact 24780

aaaggtctaa ttgatggggt aaggttcatg agcagatcag atctttacaa ggaaggttct 24840
tgctgggggg caagattcaa aacccctatt cagaaccagg agagaagaaa gtaagaatgg 24900
gagcaaatgt aggtaggttt ggtgaggatc aggaaatgga ggggaagagg tcattaaatg 24960
tggtcctggg gttgagcagc agattggaag agaatggcaa aagtttggga gctgattagt 25020
gataaggaaa aggtttgaga atccgatgaa gattagaaac catgcatttg tagtgaaact 25080
tgtttctaag attgtgccct tacccacctc cagctgtgct ctgataggtg aactatacaa 25140
ttgatgtaag gctggcagat aatcaaagcc aagaaatttc atgttttcta tctattttca 25200
cttccgtgcc atatagcttc tctcatatag tactcatatg ttttgagttt ttggtgaagc 25260
tgattaaaca tttaacaact tttttttta catttttat tgtgacaaaa catatatagc 25320
ataaaattac cattctggta aatttaccaa aacccattta tagcataaaa tttaccattt 25380
tggtcaactt actgtaaatt actgccagtt ctgtacagcc atttgtttaa tgcagtctgc 25440
tctaataccc tccaaggcaa aattgcaaat aaaaacctgt tttgatttct acttaatttt 25500
ggtaaattta ccaaaagggc aaattttatg ctataaattt tttttgtttt gacatttaaa 25560
atatgttttg taaaaaaaat tgacagataa aattgtatat ttaacatgca taatgtgatg 25620
ttttgtatta tgtataacat gatgtatatg tcatacataa catatataca ttgtagaatt 25680
gttaagtcta ggtaattaac aaatgcatta cctcacacag ttatcatttt tgtggtgaga 25740
acattttaac atccactctc tttaaatttt tcaagaataa aattttatca tctggtcatg 25800
gtggctcacg cctgtaatca caacactttg ggagggcaag gcgggaggat tgattgagac 25860
tagaagttct agaccagtct agacaacata ctgatacccc gtctctacaa aaaataataa 25920
aaaagaatgc aatattgtct aaatttcttg aatttatttt tcctacctaa ttgtcgttat 25980
atatccttta accaacgtct gcccatttcc ctctcctctc taaccagtcc agcctctgag 26040
accattgtac tttctatttc tatgagatca gccttttaag attctacatg tcagtgaggt 26100
cgtgtaagta tttgtctttc tgtacctggc ttgttttacc attgtcacca tttataaagt 26160
gtacaattta gtggcattaa gtacatacag aatgctgttc acccattact actgtcctac 26220
aattttaatt tgaggcatga ctttgtcatc tctcttctat cctttgtacc attcttgatc 26280
tataaactca acaaagtcct taactgagta gaatattttt gtatgtgtgt atagatgtgt 26340
gtgtgtgtat ataggtatgt atagatatag atatagatat agatataaaa tttttttttg 26400
agacagagtc ttgctctgtt gcccgggctg gactgagtgg catgatcatg gctcactata 26460
gcctctacca tacaggctca agcaatcctc tcacctcagc ctcccaagta gctaggacta 26520
catgcatgca ccaccatgcc tggctaatta aaaaaaaatt tttttttttct agagatgggg 26580
tctctgtgtt gcctaggctg gcctcaaacc tctaggctca agcagtcctc ctgccttggc 26640
ctcccaaggt gctgggatta taggaatgag ccaccgtgcc tggccagtat ttgtattttt 26700
gatactgagc ttaattttgg caagtgctgt gcaaggcaca aattgtggtg ttagctttct 26760
aactatttgg ttgtaattat ttgttaatat ctgtctttcc tagtcatgga aagtccatga 26820
gtgcaggggc tgtgtctgtc ttgtttacaa ctatatttgc gctggccagc acagtgcttt 26880
acacatagta gacattcagt gatgtttctg aacaaatgaa agtccttcgc tgcaagaaga 26940
atttatttca gtttaataag ttctttctga atgcttcata tgaagtaggc aaaatttgta 27000
ttcacatatt aatacagatt ataggtatgg cacattccaa tgtctattta attagtagaa 27060
atcaaaaaga gtttttattt gcaattttgc cttggagggt attagagcag actgtgcact 27120
aaacaaatag ctgtacagaa ttggcagcaa taatcccaaa gtttgaaggc tgtaggttaa 27180
ttaggataac ttgacaggaa gtgagttaat caaatttgag agtttaatct tccaataatt 27240
tatgttcaga catacttcaa gtatatcagc aggtaacagg aactttagtt gcagaatgcc 27300
ccaaaacaca agaactccag tggattttct ggcttccagg aatgttttgg aggaagaaaa 27360
accaataaaa tgatttgggg gtcattttgt tccattactc tatattaaat atactagaat 27420
ttaaaatatt aaatttttaaa agataaaaag atgcagttta cctattaaca aattaaataa 27480
tttaggaatt ctacttagtt ctgtaatact ttaatatgag taaatatggg catttctgtg 27540
ttagctagaa ttagatagag tattgccatt tttttcaact ggcttatggt taaatggaag 27600
taaaggggca aactacacat ataagaatta gtagtacaat atttaataca cccctgtaag 27660
agttatcaca gtgtgtcctc tgtgaaaagt aagggctcca tgtgtgcttg tgaaaaaggc 27720
cactggaggc ccttttcaaa aattaaatct gcctccagca aggtgttttt cgatcacatg 27780
gaaaggggaa gaagaagcta tcaggagctc tggggttttt tttgttgttg tttgtttgtt 27840
ttgccacttt taactctcaa gctaaaactg gggtttcatt tgaggaacca gtaatagaaa 27900
atttcttatg tacattcagc aaaatctagt actgagtggt tactttggct tttcattgtg 27960
gggattgtgt gtgtgtgagt acatgcacgc acttgtgtgt ttaagcgtgt aaggcagaca 28020
gacagtgggt acaggtcttt gaaatggact tcttggcaaa agtaatagag aaaaagagga 28080
atacaaataa gggaggaggg acagggaaga gcagagtcac aggaaacagt gaatgaggct 28140
gcagtctcag tcgccctttc tttgtccctc cagtgttgtt gcctgtctta tgatgatgct 28200
ggttttcagc caaccttgag tgagtaaaag ccgggtctga ggtctcagtg cctgcgtggc 28260
tgatatgagc agcttgcatt tctgactggg ccctggagca gcaacagcac agatttccag 28320
gaacagttcc tcttgtcatt tttattcctg agtcatcaaa tttagttatt cagacgtctg 28380

ctgttcccag ctacatacag atcaaacaag cagggaattt ttttctttct cttctctccc 28440
tctttttttt gtatttccat cttgttttgt ataccttttc tttgtttaag tcaagcattt 28500
gaacatcact agttaccatt tcctttagca agcataggac ttctgtctta cttaaatgtc 28560
ttctaatgct gtgatgtgtc acagttagtt gagacgttaa agatgttcca tacatgtgac 28620
tacattggta aatctcaaaa acatcatatc gaatgaaaag acaaagttgt gaaatgtctg 28680
atgtgatgcc atttccttaa aaagtcatgt aaagcaatcc taccacatct catagaatct 28740
aaaaggctat actgatgcta agatgcacta ttattttctg tacactgaga aaggggaaaa 28800
attgccactt aaattgtgat gtaatgtctc attatggatt gtaagataca tctacatttt 28860
agaaatggta taatgttaaa aatatgcatt ttaaattgaa ggtaaattta attaaattat 28920
ttcaaaggaa ataaggtaaa tgtattttat tgaagcatag ttatgtaata aaaatagaaa 28980
agcatgcata ggaatgctat ttagccaata caggatgtgg taatctctat aaagggaggg 29040
agggaaatgg agggggggagg ccagaggagg ggcctcatct ctgtagttta ttttttaaaa 29100
ttataaagca aatattacca gagttaagat tttacaaaat tcattggtaa gcacctataa 29160
ttttctgagt gctttcagta tttcataatg aaaagtatgt attttaaagg tacgttatct 29220
atttatttt atttatttt tttgagacag tgtctcattc catcgcccag gctggagtgc 29280
agtggtgtga tcttggctca ctgcaacctc cttaaagcta cattatttaa aagtcacata 29340
caaagcaagt tgcagaagcc tgtatgtagt ggattctatt ttttttaaat agtatttaat 29400
tgtatgttct tctacacttt ttcctatgtc catcttacca tagctgtgcc ttttttggtg 29460
gaagtgagga cagattgctt tccacatctc catttttgtg tctgaattaa aagatggaca 29520
agtatcatgt attatcttag tagtcatcaa acaaggaaaa aggtttcttt gtttgcttgt 29580
ttttttaga tgaagtctcc gcccaggctg gagcgcagtg gcacggtctt ggctcactgc 29640
aacctctgcc tcctgtgttc aagcagttct ctgcctcagc ctcctgagta gctgggatta 29700
taggcgcctg ccatcacgcc ggctaatttt tgtattttga gtagagacaa ggttttgcca 29760
agttggccag gctggtcttg aactcctgac ctcaggtgat ccacctgcct tggcctccca 29820
aattgttggg attacaggcg tgagccactg tgcccggcct ggaaaaagtt tttaatggta 29880
aagatgtcat ggaatgaata ggattggctg gcattatttc ttgctgttaa taagcagtga 29940
gaaatgtttc cattatatgt ttctttgaag ccagctttct ggttgctccc ttattctttc 30000
tttctcaggc atgtggtatc tagaaagggt caggagtacc ttgataaaaa ttattgtaca 30060
agttgagcaa aactcagtag tatcatgcct agagatctga taaagaggca cttttaaaat 30120
agagccttga agctaacaaa ctttttttt ttaacccctt tactgaaacc taaaaagagg 30180
tctgcagttt ttccctcctg tcaccagatg aaagggctgt agtagtgttt gcttattcct 30240
caggcagtga ggaatcaaag gactgaaggg gtgtttgttt tatcatacta cttgtaggga 30300
ccccttacct cctataccca tgaaaaagga atatttctta tatcccataa tattctttta 30360
gtatcacact taggttttaa tgtccttact gttagagtaa aataatttgg gcaagaccaa 30420
ttttttaaat ggcaaatata gtcacatcac ttgattcaga atctcactcc ctagcatctc 30480
gcgtaatgac tcataagaaa gaaaaagcta aatgcatgaa gaggttcact ataccattac 30540
tataaaaaag taaaaatttg tttttgttat tttttttttg tttttgaga tgtagtctct 30600
ctctgttgcc caggctggag tgcagtggcg caatctcggc tcactgaaac ctccgccacc 30660
tgggttcaag caattctcat gcctcagcct ccggagtagc tgggattaca ggcatgcacc 30720
accatgcctg gctaattttt gtattattag tagagacggg gtttcaccat gttggccagg 30780
ctggtctcga actcctgacc tcaggtgatc cacccacctc agcctcccaa agtgctggga 30840
ttacaggcat gagacactgc acctgggcaa aaaaagtaaa aatttgaagt caacttaaat 30900
gcccccaaat acaggaatgg ttaaatgctg atgacctact ttatgtacta tgatgcaaat 30960
gcaaatgata atggttactg tgcagcaaca tggaagaact tggtgaaagc aggatgtacc 31020
acatgactgt aactccaaaa tgtgcttgca tgtgaggaga gatcaaggga atatagaaaa 31080
atgaagacat agttgtgtta gggtggtaga attttgagtg gatttctccc cccactattg 31140
gtaaaaattt ttgtacttaa ttcgttgtgg gcagccagat cttttaaagg taaatttgaa 31200
tttctcttta agaaaatggc agacagaagg atgggggata ctagaaaact aaaagtagtg 31260
ccccctttga agataaaact aaaacatttt aagcctggaa ttgctttagc agtacatgta 31320
ttgattattt aattttgtcc tttagaagaa agttggccca acacaattac atggaagttg 31380
ggttattgaa gaggattgat aaaagaagtg ggaaggtcag gccaggtgtg gtgggtcatg 31440
cctgtaatcc cagcgctttg ggaggccgag gtgggtggat aacgaggtca ggagatcgag 31500
accatcctgg ctaacacgtt gaaaccccgt ctctactaga aatacaaaaa aaattagccc 31560
agcatggtgt tgggtgcctg tagtcccagc tacttgggag gctgaggcag gagaatggcg 31620
tgaacccggg aggcggaggt tgcagtgagc ccagatcaag ccactgcact ccatcctggg 31680
cgacagagcg agactccgtc tcaaaaaaaa aaaaaaaaaa aaaaaaagtg ggagggtcaa 31740
agccaatgtg cacgtttatt cttttgtacc accaaataaa ccgaattttt ggaccaactt 31800
tactttagtc aatatttggt tatttgcttt gaagctattt gtttgcaata atgcaattaa 31860
catcctcaac agagccaaag tctgaccctg aaatggcagc atcttaacta attttatcat 31920
ttgaatagaa attgaaccct ttgagggcat gaagtactat ttcttttctc tatacttctt 31980

gtttttggtt tttgcttcac caccgattta tgactaccaa aaccccacag atcctgtaat 32040
taactcacag ctgccatctt gccaccgtag atagaattca ggagctggtg atgaatgggg 32100
agagaataga aaaactaata gactgtaaga attttagacc tctgtggcct tgctgaacaa 32160
ttaatccagc cccttcactt tacaggtaag taaacaagtc ccagggaccc agcagtgaac 32220
ctgtgtctta ggacttcaaa tctagtgcac tttctgttat acttcaagga gagaactgga 32280
gggagagcga agccaaaata ctgacatttt agaggctgct ttttaagaaa ggataggaca 32340
ttggacttgg tgtcattatc atgttaagtt atagaatttt taaaaacact accactccaa 32400
aacaaaaata ggtaaaagat acgaacaggc agttcaccaa agaagaaata cagtaggcat 32460
gaaagaatgc acaacataat ttaagaaatg taaataacaa atatctattg ctatctcaag 32520
agaaaacatt tactatgagc tttattattc aaaccattgg cctcataggg cttcttgttg 32580
acaatctacc ctaattttag gaaatgtaat tataacataa aaaaatgaat gcaaatgata 32640
aaagacgaag tataacccgc tcaataaatt ttctgaaagt tatagtagtt ttaaacctct 32700
ttttattcat tccctccagt tctgtctagc acctgtaaag atgaattatt ggctgggtgt 32760
ggtggctcac gcctgtaatc ccagaacttt gggaggccga ggcgggcgga tcacgaggtg 32820
aagagatcaa gaccatcctg gccaacatgg tgaaaccctg tctctactaa aaaatccaaa 32880
aattagctgg gcgtgctggt gcatgcctgt agtcccagct acttgggagg ctgaggcagg 32940
agaattcctt gaaactggga ggtggaggtt gcagtgagcc aagatcgcac cactgcaccc 33000
cagcctggca acagagtgag cctccgtctc aaaaaaaaaa aaaaaaaaaa aggtgaatta 33060
tctcaggttt tatatggcat tgatggacca ctattgaagg aagctgtgaa ggactcaaag 33120
cagttgagaa taggtatcgc cacctaattt tacttgcctt tcctgccaca cacaaccaga 33180
gaagctggtt catctcttac tgaggaaata ttttcatgct tattcagaag attttgcagg 33240
atcctcagga agacaaagcc aaagaacaag aacagtcaaa gtaggaacag gaccaaaagt 33300
ctgcacagtc actggaaggt gtcatgctga agaaggcagg gatgggactt tgaaatgagg 33360
ccaagtgcat ttcagtaact gagtgggtta tctgttgttg gcaagcaatg tgccatacct 33420
ttgaagggct aagctagccc aggagttctg ataagagctt tgaacttaat atcgctgaca 33480
caaaaatagg tctcctcaga tcctatttgg aggtaaaccg ggtaaaagtg ataaatagta 33540
ttgcttttta aagttcatta catattttag aggtgagtca gtcagaccat atagagatgc 33600
ttcattttaa tcctctgtag caagaattgg caaaaatatt ctctgtaaag agccatttag 33660
taaatatttt agagtaggct gcatagtctc ttttgcagct actcagcttc atacaacaag 33720
taaacaaatg agtgtgactg tattttagta aaactttgtt agacgctaac attggaatct 33780
tacacagttt ttatatgttt tgacatattc ttttgatttt ttccaaccat ttaaaaatgt 33840
aaaaagcatt tgtagcttgt gaggtgtaca aaaatgagca gtggaccatg ttttgcctgc 33900
agggcacatt gtgctgatct ttactctata gcatcactgc cagtagaaat ataatgctag 33960
ccatatgtta tggactgaat gtttgtaacc cccaaaatcc atattttgaa gctcaaaccc 34020
ctggtgtggc aatatttgga gatgggacct tgacagatgt acttaaggtt aaatgaggtc 34080
acaagggtgg ggtcctggtc agataggatt agtgtcctta caagaagaaa caccagatac 34140
actcattgag aaaatgcaga acattgtaac cttactggtt tagaaaccat aatactggca 34200
ggcttgctgg tcttcatttt tattattttg aatgtcttct ttcaagtatt ttatgtatcc 34260
agacccaatt tgtatgataa gggaatttct ttcaaactat taggtgaagt ctttaattat 34320
gaagttggct taagccatta agtagtatta atggggacat ccatcttaga aattaactgg 34380
aaatgtagac tagaaatata aaaaagtgtg gtcatgacta atgtctgtat ccatttccca 34440
aaaaagagat ttgcgtatgt cctcacattg cagaaccaca ctgacactca gggaaacatg 34500
ctgacgtcat tcttccaaca tccttagttt ggatttcatg aaacattttt ttccatttct 34560
ccttttttctg tgtttagtaa tcctttcttg gattcttgaa atcagaggca cttattggag 34620
ttgataaatg gcagttcttc atgtcactgt ttagaaattt aacttagctt atggactagt 34680
tcaggcttag ataactctgg aagtctttgc actttcaaat atacttttttt atagttgagt 34740
ttctacacaa tataacattt tataacttaa atggaagtaa agtgaaatgc caaaatgcca 34800
gttgattctg ttaatgtgtg catccctcca ttccagaatt ccaaaaatat cagacataat 34860
gacttttatt tgtattttat attattagat ttatcagata taaaattgtt cttttaaact 34920
gctatttcta aacactcttt gtttctttag ttatcttgtg aactggtaat agttttcaga 34980
ctggtgcctg tccatgaacc actttcagca ttggtccaga tgactttcac ttataccatg 35040
agccagcaga tactgtagat tatttggctg tttttaaaga gaacatcata tcatagtgtt 35100
ggttattaat atttgagcaa gatattaggg ttgtctcgat gtcccaatca ttgtggagtg 35160
cttagtgtgt gtttgttttt taagaattga cttggaaaca tttatcttaa ggtgttaggg 35220
ttatttttat tctttagtac taacggtata aactgaattc cactgaaata gagcttattt 35280
tgatgtataa tttttgaagt tatttttttat aagtctgtgg gtagatagca gccaagtagg 35340
cactttattt ccctaataga aaaaaatata tatttttgtc aaatatttct gttttattca 35400
ttcattcaaa gtatatttgg aatgtttatt tccaggaaat tttggaatat acaatacaac 35460
cagcttctta taactccact ttaagtgagc cataggtcaa ataatgacca gcaaaatgta 35520
atgacacgtg tgcctcttac tccctgttgg aggaattgag gcactctggt aaccctgtag 35580

gcctggatta gtccagttca ttggcagcag cattatccag attttattgt ggccggcaac 35640
ggtggctcac accggtaatc ccggcacttt gggggggctga gttgggcctg ttgcttgagc 35700
ccaggagttc aagaccagcc tgggcaacat aggaaaccct gtctctacaa aatatataaa 35760
aattagctgg gcgtggtggc gtgtgcctgt agtaccagct acttcggagg ctgaggcagg 35820
aggatcacct gagcccagaa agttgaggct gtggtcagct ctgattatgc cactgcaccc 35880
cagcttgggt gatacagtga gaccctgtct taaacaaaca aaagagattg tattgtgttt 35940
tgaaaaacat agtttttaga gaaattctgt gttattcttc tatttccatt acagcttggt 36000
ttaaagaaac tatggaaagc tacagactaa cctccttttc ctttggtttt ttttccatgt 36060
tatagcagag aaccagccag ggccagattc ctgggcttgt gtcctctgca gttgcacagg 36120
gctctgctca gaagggtccc atacttggct taatgctctg ctgtcagcat cttgaaaatt 36180
tttaataatt ttatatttca tcttgttttg taagttataa gcccagtggg ccagtggagg 36240
ctgttctagg ggtctgggtt tggctcatga gaggtacatc cttcccacct ccccaggatg 36300
ggctctcagc agctggcttt ctgtaccctg gcaccccagg cccttcttgg cccctacccc 36360
gtgttcactg ttgccctcac cttgggtgat gactgggtca catgggagag gaggaaaccc 36420
atgttctgct gtcaccctcc acctcttgct gggtcctggg tacaagctca gggagggttg 36480
ggatcaggcg tctgtgactc acagtgttgt ggggagtgat ggtggtcatt ttatcccagg 36540
ctggcagcgt gctggcatat tgatgagcag ctgctggcat gatgatttgt cctgaccctg 36600
gaccggactc tgtatcttgg ctttagtcta gtccacttct cagcattcct tggaagccca 36660
cctaaactgc agagacaaag aagaatatgc tttgggattt taaagatgga accccatatg 36720
ctatcaattg gaaaggagtt cagattctgg ctgggctgaa tcactgattt tccaggcaag 36780
tggaacttta atgacccact agtgcccttt gttcactgtg ggctggctgg ggaggaaccc 36840
ttgcctttaa agaaccctgt gctctttaaa gtttccaggg cacaattgag atccttgagg 36900
tacagttcat ccatgcagac ctagcacctg ctttaggtaa gggctcttgt tacttcttgt 36960
tcttatcggg ttcttatgtt cagctcagct tcctgcttcc tcagagagaa gtggtaggtt 37020
ctgttttagg cttttctttt ttggtggtgg tgggaggtgt caaaccgcat tattgagggg 37080
agtgggggtg gtaaagcagt ggtgaggacg atgagaagcg acctccttag ccaggccagg 37140
agctcacctc ccagaccctg cccttcaggc actgaaaact gttttaggct tttctgcatg 37200
gttggagttt actcttagtt gctgacagtc actcatcccc tggagagggg tactgagcat 37260
aatcactgtt ggtattttat ttactaccag ttcttcatca ttaccgagac tcattccatt 37320
tggaaaataa acattgcctc atctgttcat ggtatatttt tacacttgcc tttcactcac 37380
atgttgcata ttaactgaaa agatcagcat gatcttggag cctaatgtga tctggatctt 37440
attaccatgt atgtttttat cgtcaggtgc tgctttggac tctggcccat tcctcccacc 37500
tgtcttcatc ctgtcggggt tgatctctgc ctggttccat agtcacttgg cttttggagt 37560
tttatttctt ctctaaccca ctcttccctt ccctgtcaaa tataattctc tcatgagcaa 37620
ctgtgcctga taatcccagt gtataggcca tagctcaggt tgttttaatg cttgcaagga 37680
agtctcttca gattactaca gggtaaaaga atccattccc tttcaacctc tcattactta 37740
ctgaatcaat tatttattta cttgtgtaat tcacatccat gctcaggtat gtgtttagtt 37800
tacctctgag actgtggctc cctctagaca agggtctttg tctggtatta gtccacttcc 37860
tcagtgtcta atttcatgag gacagaatca gggcaagacc tactggctaa gacttgctgc 37920
tgcctcagtt atttggttta ctgtttaagc ccttgagcca ggaggtgaca gtatcttttg 37980
tgttccaagt cagataacca gtaacagtaa tattctgatg taggtctaag gggcaatagg 38040
agcctgaatc tgagcccctt ggcagggatg gtttccctgg gttactggat ttgagtcttt 38100
gtcttccgta atagtaactt ctgtgaccct tggctcaagg agtcctttct aacctaaatg 38160
ccctcctgaa gaaagtgctg atattcatca gaaaaaaaag tagggtttgt gttatgttac 38220
catctgggag acattggctt taatcctctc ccctcttctt ttccttttaa aattttcaag 38280
aatccccaga attgagggtc aatcctggat cagaaggggc agaaaagtca gccagtcaga 38340
ggtgagaaag gggtcttgtg ttccagttgc tctactttag agggcttgag ggccgcatta 38400
ggagcaaacc aaatactctg aggtcctagg agaagattgc tctaaatcca ggatgatttc 38460
atcatatgac ttgagatgga ggggttcttg attagtaatg ggctttttga ttggcaggaa 38520
gtaggtaaac tgcaggcaat aggattggcc atagaaatat agggaactct atgctttcaa 38580
ctcttgtgaa gttaggacca aaccaaatta tttggggttc attggtggct aagaaattaa 38640
cttttatact tatagttaaa aaaaaaaaac ttaaagaact tttattaatt cttccccaat 38700
agagtgattt ttttttttt ttaagagacg gagtctggct ctgttgctca ggctggagtg 38760
cagtggtgca atcatggcca cggctcactg cagccttgac ctcccaggct caagtgaccc 38820
tcccacctca gcctcctgaa gtagctagga ctacaggcgt gcaccaccac acccagctaa 38880
tttttgtatt ttttgcagag atgagatttt gtcatgttgc ccaggatagt ctccaactcc 38940
tgggctcggg ccatctgccc gccttggcct tccaaagtgc tgggattaca ggcatgagac 39000
accatgccca gccagaatga ttttcctttg ttgaattctc aaaatttttat tctggcttag 39060
gtaattgagt aggaatggca gagagtaata attttcagtg taatctttag cattacataa 39120
acctctacaa acatgaagta aaaagttctt tacagacttt ctataggaaa aataaatgtt 39180

tatagatctg cttatcaact gtttgttctt cctacctctg cattttcatc tattgataac 39240
tgtcagagtt caaggtcagt taattgtaca ttttctggga gtttttctca ctgttaatta 39300
gtaagacttt gaatcattaa gtgctgatga ctacatggca caattcatta aagacagtca 39360
tactttaacc ttagtacagc tataatggga cggtgaaata caactaggaa agtatcttat 39420
taaagtattt tacattttac taataaaatt tatttaaacc ctgaattagt gtttatttt 39480
gtttccgaat tctcagtaac atctcagtag atccgtgaat cctgccaaaa gacgtatttt 39540
aatccaaaga tctttctgca tctataattt accacactaa agctcacatt atcattaaaa 39600
gcagcattat cattaaattg tagacattta ttagttcatc tccaggccta atgggagtgg 39660
tttggagtaa acctttgtag aaacaataat atgttttgat aaatgcatat gggcgtggag 39720
tggttatcac ttacctcctc atgaaatttc tggaaaggtg atccggaaaa ccaggacaca 39780
tttatggtaa gatataacac cttaccataa cgattaaggc tcattgaaaa actctgactt 39840
aatgagttta ccaaaaaact gttctatcca catctcatca aaccgcccct gaaattccct 39900
tgcctctgtt aaatttttct agacagtctg aatagaggca tgtaattttt ttggattttt 39960
ctgtggttaa ataaatatcc tttacaactc tctttattct tgaatatcca taagagttta 40020
tatttatact gtatgtttgt tattaggatt cctttcaatt gctatataaa aaatgtaaag 40080
tctgtttact gccttaaacc ttctggtgta tttttatata aagtaacacc cttaattcta 40140
acttggccaa caggtaggat ggtattatta ttatcttcat tgtacagata aggaaactga 40200
ggctcagatt gactagatca aacaggagtt ttctggaaaa cctaggacac aagcctaaat 40260
ctttgaactc aaatactgct ctacactgaa ttacagttat atactgattt ctgttgtaaa 40320
ttcttagaga agacagacat agaaattagt aacttgagtc agtagcggct ttgttcaaac 40380
acaggcacat gcatatttta tggtatatgt ttatatctgt gtaatactca tcataaatgt 40440
cagatttata atcgatagtg cccatttcta aatttatagt tgaaacctct gataggaatt 40500
aggaattatc ttaaagtcct agaataataa atattaagat tttgaagact gcttaaaaca 40560
gtgttccagg ctggattttt tccttagttt ttttttcctt agtttttac cttagttttt 40620
tccatgtaat gaatgaggag ccattggggt tggtgtttgg ttgtaaaggg caggttgact 40680
tcaccaggag gtgtttcttg gtatttatgg atcctctttt tcacactaat cctttgatta 40740
gcttcttctt atagttcatg cttgtacagt tgcagctgaa tggtaagtag gtagaaatat 40800
gcattgacta atgttgaact atatctagga gcgtcatatt catgctacta ctgagcactg 40860
tgacctaggt gaagtgtgag attaaagagc tttattgctt accacgtttc tttttgatct 40920
agttagagtc atgtgaggca gtccatgact tttaggatgt tataatataa tccacatagt 40980
actttatact tttcaatttg tttttatata tttttcctta ctataagcct gtcaagtcta 41040
gtgagatgac atggttatta atatattaca aatgaagaac tctgacttgg aggtagtata 41100
gcactgtggt taataataga agttcttata tttttatatg aattatatat tatgaaccag 41160
accagaatac tagctctacc aattacctag ctgagtgatg gtgcagcaac ttaatctctc 41220
aaagccctgg ttttctcatt gtaaaatgga ggaccttata ggattgttgg gtagcttaaa 41280
tgagataata ggtagaatta aaagtaacgg caaaaaccgc aattactttt gcaccaacct 41340
gatatttaga acagtggagg catatagtag gcattcaata aatattggta gtattatttt 41400
cagaggttat tactcattta actcaacaaa aattgagagc ccctttctat acgctagtca 41460
ttgtgctagg catcaggagc acagggcaaa cctggtaggg tgcttactct gatggaattt 41520
acattttagt ggtagaaatg gtaaatgagt caacaagaat tttatggagt gacaaggact 41580
ttgaggaaag taaaacggga caatgcttga gaagggtgct agttgggaag tcctatcaca 41640
cgaagtggta tttgatctga gacctgagtg accatagcca catgaatacc tgggggaatc 41700
acttcccagg aggaaacaga agcacagaag ccgaaaggca ggatcacagg cggcacattt 41760
tattggtgca gtgtggaggt agagggactg gattgtgtag ggtttgtagg ccagggtaaa 41820
gagcttggat ttcttctgag tgtaggattt tgagcagggg agtgatgtgt cttattttgg 41880
gcgggcgcta ggccaggtac tgtgaaggaa ataatggtat agaagaaata tgtctttctt 41940
gctagaatct tatgtgacat acaactagtg gtggctgatc agcaattaga actgcagtgt 42000
tcttgctggt gtactcctaa tctgatcatt ggatgatatt tatcacgtat ttttgcagca 42060
gtcatctcta aagtggcatt atggttctag ttatctatca ttggatctgt ttgctctaac 42120
agtccgaagg ggcatgaagg gtggaaacag agtagaggga aatgggtgaa aaggctattt 42180
ctaaacacta gacacaaaat agggattagg aagtagtggg ggttaggaga aaaagcaaaa 42240
tcctgttgtg catggagagg gacattgttg aaactgtggc ttcttgagat tacatgagtg 42300
aacctgcgtg gactgggtcg tcttccctat tcaccaagct gaatccaggc ccaatccaag 42360
attactgctg cctttgctct gaggctttag actgttagta gcttttctcg tctctactga 42420
ggccaaaagg ggcagtgata cccttgaatt ttcttcttaa aacagggttt catttccttg 42480
gaagtttgtt tctttgaatc tttctgtcag tttaactgtt atcatcaatt ggttagcatt 42540
ctaataataa ttataattat agtaaacatt tattgagtgc ttacgaagag ccagttccaa 42600
gctttttat ctccattatt ctgctacttt ccttctcatt ttacagatga ggaaaatgag 42660
gcacagagtg gttaattaat ctgtttgagg tcccgtagca ggtcagtgat gccagggttc 42720
aaacctacac ttaactctac actagagact gttttcttaa ttatttcttc acaatcatat 42780

gtttaatgat tacttattga ttatttagtg gtctgataag aagagggagc ggtgctcttc 42840
tgttggagaa gaaaggctgg ctgatcaaga cacactggtt ggtttgaaga aaaaatatag 42900
atgttaattc cataacacca cactctaaac atttctactg gacgagttcc acctgtgtgc 42960
cactcgaagt cggatgcagt aaggaaggct ttttattgag gagagaacga atacctctgt 43020
attcaaaaga gagtgtgttg ttccttatag aagatggaag ggggcttgcc agtgacagat 43080
tatgatgatt acctccttag tggtttttt ttattgcaca gactataata ataattataa 43140
aaatttgtaa atcaagaaaa atctttaccc caacaaattg tttttatatt tttcatattc 43200
atatttttc atattcattt tatttttca tattcaaaaa taaaacaatt ttcatattca 43260
tttcaattgt ttatttttc atattcaaaa ataaaacaat tttcatattc atttcaattg 43320
ttttattttt tcatattttt catatttcat atttttcat ttttatacac agttatagag 43380
tatctactta tttttcattt aacattatgc tataacaggt tttaaagtta ggcagtaagt 43440
ctttccagtg aaaaggaaaa gactgaaaca tgaaaaggtt acgaggcatt actaattgat 43500
tggaatgatg cctaccaggc agcccatgtc agttttggag ctggccttgg aaggaacagc 43560
tgtgtatttg gaccttggag cagtctgtca ggcttccaag aaagtgggca aagaaggaag 43620
gggggttggg ggcacaagtg gaaaagtaca gtggtttgga actgatgtga aactatcaag 43680
ctggttgagt tcagtgttag aaaataggaa aaacaaaaca atttctacct gatatggtac 43740
tttaattgag gttagcactt cagtgaatag gcaatgcatt ataatttgtc agaattatca 43800
atattttgtt cttgtcagtg gctttatatt taatagttac atgttataaa gccagtttta 43860
tcagtgaaag aacaaagttc ttgacaaatg ctattttagt gataaagctg ttattctgta 43920
tttaaattca gtttaactca aagtggtttt taagttttac atttgtataa ctgtagacta 43980
gccatatggc attcaaaggc ctccaagata taacttgaga cactcttagg atggaatccc 44040
atattcagaa tctttactga cagctcccca tatccaaatt gccttttctc tcttgggcct 44100
ctgggtgtga caagggacaa cttgagtgaa aggccacagg acaaagtaga gatgctgtta 44160
aatctcaatg cagaaggcct gagacaactg atagagacaa ttgtttctct tttctccctc 44220
ataaaataag taaaattcaa agtgattttt ttttaatttt ttgctttgac acattgtgtt 44280
tagtctgatt gaggtcatct tacacaacag aatccaaaat cctgagaaaa taatttatac 44340
aactagaatt ccaaaattag gttttgcatt aagttcaagc tatcatttct ttaagaaaac 44400
ctacactgga gttaccaact ggaatatttg gttatgagac acagtcatca tgtcagagta 44460
tgtagtttgc ttctaacact tatttgtgtg cttttaagac ctgctccctg ccacacttta 44520
tcataacaaa aaacaagaga cacatgttag tactgtagca tgtatagtca tggggaatgt 44580
gttctaaacg aaatcccatc acagtggagt caacacaata tttctagggg agaaggtctt 44640
cctgttaaac tcatgtgaag gcattcttct ctactatggg aacttgttta tgtgccctct 44700
agaagacagc tgagatggtc ttcagtgaat ctgttcactg acatgttgct agtttcttta 44760
tgtttacatg taaacagtgc ttagaagcat ctttccccaa gctgttcttt ttttgatgga 44820
ctcccccttt ttgggggagt atcagttgaa tgagcactgt tgtactctca gcactagaca 44880
ttctgttatg tgtgtgtgga gcttccttgt ggcttggtag cattacctct gcattacatg 44940
ctgcagcctc accaggcaga agggccattc acttctgctg ctacttacca gcacctctgt 45000
ttcctcgagt gtgttaccct cccaccacct gggccttgga cctgggtccg ttgggctatt 45060
atccatgctc ccacctgccc atctctgctc tgatataatc tactaggata ataaagtagc 45120
ttttccctac cataatgttt tttaaataga catacagcaa tgtataatgg ttaagagcat 45180
gatccaaact gtgtctgatt ttgaatcctg gctgagttac tagctatgta atcttgagaa 45240
aactactcaa tctctctgtg ccttaatttt cttaggtata aagcagatac taattatgcc 45300
atcatagggt aggtatgagg attaaatgag tgagtatttg taaaacactt aaaacagtga 45360
ctgacttggg ctaccccttt tgggtcccct ccctttgtat gggagctctg ttttcactct 45420
attaaatctt gcaacttcac actcttccag tctgtgtttg ttatggctca agctgagctt 45480
tcgctcgctg tccaccactg ctgtttgctg ccatcgcaga cccgccgctg acttccaccc 45540
ctctggatcc ggcagggtgt ccactgcacc tctggtccag cgaggtggtg cccattgccg 45600
ctcccaatcg ggctaggggc ttgccattgt tcctgcacgg gctaagtgcc ctgggttcat 45660
gctaattgag ctgaatagag ctgtaacact cactgtatgg cccaaggttc cattccttgg 45720
aatctgtgag gccaagaacc ccaggtcaga gaagaagagg cttgccgcca tcttggaagc 45780
agcccgccac catcttggga gctctaagaa caaggacccc ccggtaacat tttggccacc 45840
atgaagggac ttccaaagcg gtgagtaata tgggaccact tttgcttgct gttctgccct 45900
attcttcatt ggaattggag gaaaataccg ggcacctgtc agctggttaa aaacaattag 45960
catggccacc agacttaaga ctcaggtgtg aggctgtctg gggaagggct gtctaacagc 46020
ccccaaccct tctgggttgg gagcgttggg ctgtctggaa ccagcttcca ctttcagttt 46080
tcctggggaa gctgagggcc gactagaggc agaaagctgt tgtcctgaac tcctggtgtt 46140
agctggttga gatcatggcg cagccagaag tctctactca acagtcagcc atgcgtgcac 46200
ccctaccttt ccttctgacc tatacctctt ggtcctgacc acaacttgct tgaaagtgta 46260
gccccaaaat tctccttacc tctgaatcta cttcctccaa tccctgcctc ctaggtacta 46320
atggttcaga ccttcatttc ctctagcaag ctgtatctcc aaagggatct aaggaagctc 46380

tatgctgtgt ccttaagccc ctaggctctg aacccagata gtcttgtccc tggtgtccct 46440
cccaatttag gcatacagct ctcaacatgg gcagttatgt aggacctgtt ccccaccatc 46500
cttgccaggg tcccaggttt gtaaagggct aggagaagag agagagagag agagacagag 46560
acagagggga gagaaagaga gagagacgaa gagggagtca aagagaaaaa gaaagagaaa 46620
gatagaagta gtaaagaaaa aatagtgtgc cctattcctt taaaagccag agtaaattta 46680
aaacctataa ttgataattg aaggtcttct cctgaccctg taacactcca ataccacttt 46740
attgtcagtg taaacaaggg ggtagcccaa aaacactgag accactgaca acctatcaaa 46800
atcctatcaa aaatccgtaa cccagtaaca cgtggatggg ccaaaggcat tcagtcggta 46860
gcggcaactg ctttgctaaa agtagaaaaa taactttaga ggaaacctca ttgtgagtac 46920
acctcaccag ttcaaaacta tcctaagtca aaaaagcaaa aaggtaactt actaactcga 46980
aaatcttaaa atatggggct cttctgttag aataaaggta acttattaac cactgaaaat 47040
tcccttaacc cagcagattt cctaacaggg gatttaaatc ttaattacca tacaaaggtc 47100
cgacagatct aggaggaact cccttcagga caggatgata gatggttcct cccaggtaat 47160
tgagggaaaa aaccacaatg ggtatttagt aattgatagg gaaactcttg tagaagcaga 47220
gttaggaaaa ttgcctaata attggtctgc tcaaacatgc gagctatttg cactcagcca 47280
agccttaaag tgtttacaga atcaaaaaaa ctcaatctca atcctgactc aaaaggttac 47340
ctacaccctc tctgaaatga atttgcataa gaactgttgt ttatgggaat gcatcttgat 47400
ggggcaactg ggttgttatg aaatactcag gaacccagcc cagctccagg actcacccct 47460
gagcaaaaaa gcaatgttgg gcactctggt aaaggaccac tagaatccag cagcctggac 47520
ccctttcttt gtggtcaaga aaggcaggaa aaggggtgca ggactgctac atcagtgagt 47580
gcaactaatc tgataagcag aggtccatgg gtggttacac accctggaaa ggaataaaca 47640
ttaggaccat agagaacgct ctaggactaa tgctcattgg aaaatgacta ggggtgctgg 47700
catccctatg tccttttttc agataggaaa tgttcccccc aaggcaaaaa tgcccgtaag 47760
atatattctg gagaattggg accaatttga ctctcagatg ctaagaaaga aatgacttac 47820
attcctctgc agtaccacct ggccatgatg tcctcttcaa gggggagaaa cctggcctcc 47880
tgagggaagt ataaattata acaccatctt acaactagac ctcttttgta gaaaagaagg 47940
caaatggagt gaagtgccat atgtacaaac tttcttttca ttaagagaca acttgcaatt 48000
atgtaaaaag tatgatttat gccctacagg aagccctcag agtctacctc cctaacctga 48060
tgtccccctg actccttccc caactaataa ggacccccct ttcaacccaa acagtccaaa 48120
aggacataga caaaggagta aacaatgaac caaagagtgc caatattccc tggttatgca 48180
ccctccaagc ggtgggagaa gaattcggcc cagccagagt gcatgtacct ttttctctct 48240
cacacttgaa gcaaattaaa atagacctag gtaaattctc agataaccct gatggctata 48300
ttgatgtttt acaaggatta ggacaatcct ttgatctgac atggagagat ataatattac 48360
tgctaaatca gacgctaacc tcaaatgaga gaagtgctgc cataactgga gcccgagagt 48420
ttggcaatct ctggtatctc agtcaggtca atgataggat gacaacggag gaaagagaac 48480
gattccccac agggcagcag gcagttccca gtgtagctcc tcattgggac acagaatcag 48540
aagatggaga ttggtgccgc agacatttgc taacttgcgt gctagaagga ctaaggaaaa 48600
ctaggaagaa gcctatgaat tattcaatga tgtccactat aacacaggga aagaaagaaa 48660
atcctaccac ctttctggag agactaaggg aggcattgac aaagcatatc tctctgtcac 48720
ctgactctat tgaaggccaa ctaatcttaa aggaaaagtt tatcactcag tcagctgcag 48780
atattagaaa aaaacttcca aagtccgcct taggcccgga gcaaaagtta gaaaccctac 48840
tgaacttggc aacctcggtt ttttataata gagatcagga ggagcaggca gaatgggaca 48900
aatgggataa aaaaaaaaag gccactgctt tagtcatggc cctcaggcaa gcggactttg 48960
gaggctctgg aaaagggaaa agctgggcaa atagaatgcc taatagggct tgcttccagt 49020
gcggtctcaa ggacacttta aaaaagatta tccaaataga aataagccac tcccttgtcc 49080
atgcccctta tatcaaggga atcactgtaa ggcccactgc cccaggggac gtaggtcctc 49140
tgagtcagaa gccactaacc agatgatcca gcagcaggac tgagggtgcc tggggcaagc 49200
accagcccat gccatcaccc tcacagagcc ctgggtatgc ttgaccattg agggccagga 49260
ggctaactgt ctcctggaca ctggtgtggc cttctcagtc ttactctcct gtcccggaca 49320
actggcctcc atatctgtca ctatcccagg acagccagtc actagatact tctcccagcc 49380
actaagttgt gactggggaa ctttactgtt ttcacatgct tttctaattg tacctgaaag 49440
ccccactccc ttgttaggga gagacattct agcaaaagca ggggccatta tacacctgaa 49500
cataggagaa ggaacaccca tttgttgtcc cctgctggag gaaggaatta atcctgaagt 49560
ctgggcaaca gaaggacaat acggatgagc aaagaatgcc catcttgttc aagttaaact 49620
aaaggattct gcctcctttc cctaccaaag gcagtacccc cttagaccgg aggcccacca 49680
aggactccaa aagattgtta aggacctaaa agcccaaggc ctagtaaaag catgcagtag 49740
cccctgcagt actccaactt tacgagtaca gaaacccaac agacagtgga ggttagtgca 49800
agatctcagg attatcaatg aggccattgt ccctctatac ccagctgtac ctaatcctta 49860
tattccgctt tcccaaatac tagaggaagc aaagtggttt acagtcctgg accttaagga 49920
tgccttttc tgcactccta tacatgctga ctctcaattc ttgtttgcct ttgaagatcc 49980

ttcgaaccca acatctcaac tcacctggac tgttttaccc caaggattca gggatagccc 50040
ccatctattt ggccaggcat tagcccaaga cttgagccag ttctcatacc tggatattct 50100
tgtcctttgg tatgcggatg atttactttt agccgcccgt tcagaaacct tgtgccatca 50160
agccacccaa gtgctcttaa atttcctcgc cacctgtggc tacaaggttt ccaaaccaaa 50220
ggctcagctc tgctcacagc agagggctat ttatccctaa atacttaggg ctaaaattat 50280
ccaaaggcac cagggccctc agtgaggaat gtatccagcc tatactggct tatccttatc 50340
ccaaaaccct aaaacaacta agaaggttcc ttggcataat aggcataaca ggcataacag 50400
gtttctgctg aatatggatt cccaagtacg gcaaaatagc cagaccatta tatacactaa 50460
ttaaggaaac tcagaaagcc aatacccatt tagtaagatg gacacctgaa gcagaggcag 50520
ctttccaggc cgtaaagaac accctaaccc aagccccagt gttaagcttg ccagcggggc 50580
aagacttttc tttctatgtc acagaaaaaa taggaatagc tctaggagtc cttacacagg 50640
tccgagggac cagcttgcaa cccatggcat acctgagtaa ggaaattgat gtagtggcaa 50700
agggttggcc tcattgttta cgggtagtgg cggcagtagc agtcttagta tctgaagcag 50760
ttaaaataat acaaggaaga gatcttactg tgtgaacctc tcatgatgtg aaccgcatac 50820
tcactgctaa agaagactgg tggctgtcag acaactgttt gcttaaatat caggctctat 50880
tacttgaagg gccagtgctg tgactgcgca cttgtgcaac tcttaaccca gcgacatttc 50940
ttccagacaa tgaagaaaag atagaacaga actgtcaaca agtaattgct caaacctacg 51000
ccgcttgagg ggaccttcta gtggttccct tgactgatcc caacctcaac ttgtatactg 51060
atgaaagttc ctttgtagaa aaaggacttc gaaaagcaga gtgtgtagtg gtcagtgata 51120
atggaatact tgaaagtaat cctctgactc caggaactag tgctcagctg gcagaactaa 51180
tagccctcac tcaggcacta gaattaggag aaggaaaaag ggcaaatata tatacagact 51240
ctaagtatgc ttacctagtc ctccatgctc acgcagcaat atggagagaa agggaattcc 51300
taacttctga gggaacccct atcaaacatc aggaagccat taagaaacta ttattggctg 51360
cacagaaatc taaagaagtg gcagtcttac actgctgtaa gaaaggacag agaaataaaa 51420
gggaaccgcc gagtggatat tgaagccgaa agagccacaa ggcgggaccc tccattagaa 51480
atgcttatag aagaaccgct agtatggggt aatcccctcc aagaaaccaa gccccagtac 51540
tcagaagaag aaatagaatg ggaaacctca tgaggacgta gtttcctcct caggatggct 51600
agccaccaaa gaaggaaaaa tacttttgcc tgcagctaac caatggaaat tacttaaaac 51660
ccttcactta ggcattgata gcacccatca gatggccaaa tcattattta ctggaccagg 51720
ccttttcaaa actatgaagc agatagtcag agcctgtgaa gtgtgccaaa aaataatccc 51780
ctgcacttca ggccatgcat ttcaatccct gaatctttaa cctccttgtt aagtttgtct 51840
cttacagaat tgaagctgta aagctacaaa tggttcttca aatggatccc cagatgcagt 51900
ctatgactca aatctaccgc ggacccttgg accggcctgc tagtccatgc ttcgatgttg 51960
atgatatcaa aggcacccct cccgaggaaa tctcaagtgc atgaccctta gttgcaccag 52020
ttcagcagga agcagttaga gcggccgttg gccaacctcc ccaatagtac ttgggttttc 52080
ctgttgagag gggttgctga gagacaggac tagctggatt tcctaggccg actaagaatc 52140
cctaagccta gctgggaagg tgactgcatc caccttttaaa cacggggctt gcaacgtagc 52200
tcacacccga ccaatgaggt agtaaagaga gctcactaaa atgctaatta ggcaaaaaca 52260
ggaagtaaag aaatagccaa tcatctatca cctgagagca caggggggagg gacaatgatc 52320
aggatataaa cccaggcgtt ctagccggca acggctaccc tctttgggta ccctcccttt 52380
gtatgggagc tctgttttca ctctattaaa tcttgcaact gcacaaaaac caaaccaaac 52440
caaaccaaac aaacaaaaaa acagtgactg acttatggta aacattatat aagcataaag 52500
taaaccaaat actttttttc taattataaa agtcctacac taacattgca gaaaacttga 52560
ggaattcaga aaagttatta cttagtaaga gttggaatga ataaataagt gggtagttag 52620
gatggcaggc atgtgtttta ggcagagaga tacaagataa agaactaaaa ctagaatctg 52680
gtctttgaac ccctggcctg attgtcttat tcatcatgat gatttgccta tttttccaat 52740
ttctaaatca ttcctctgct gttgacaaag caataaattg ttatatttga taagtgaatc 52800
ttcagagaac tggccttgag ccagctctac aactaaccag ctctgtggcc ctttggagaa 52860
tttcttaata tttgtaaacc tcagctttcc taccagtgaa atgaagttag tcctccctgt 52920
cctgcagggt tgctgcaagg atttaacaac atgtatatgt acaaaccact tagtcctgtg 52980
cttggcctat ttggtgcttt tttttttttct ttttttttaag acagggtctt gcttgaatct 53040
tgctgaggct ggattcaaac tccggggctc aagtgatcct cctgcctcag ccttccaagt 53100
agctgggact acaggcctgc accactgtgc ctggtggcag tgctcgttga atgttctttt 53160
tccttagttc cttcctagct cttctgacag ttttggggct tatgtatata agaaggactt 53220
ggttgcctca gggagagagg atgcagtaga gttacatagc tcacctcaca tcctccaaaa 53280
gctgaattca taagtaaaca aagtgagcat ttcacccata ctttacacaa agtctagaat 53340
atttatggtg tccatcaggc tcacatactg tgaccttctg agatactttt ccctctccat 53400
tccctttcct tctccctgct ggctttttttt tttcttcttc ttctttttttt ttttttttta 53460
ctgtgaaaaa caacctatat acagaatagt acaaaaacat acctgtatag tttgaagagt 53520
aattattaac agtcttatta agaaacaatg ctccatccat gttactgcaa aagacatgac 53580

cttattcttt ttcttcttaa ttttttttctt tttctttctt tattttggcc ctttttcaga 53640
tctagacctg cagagatctt gttctttttt tgagacagca tctcgctctg tcaccaggct 53700
ggagttcagt ggcgtgatct cggctcactg caccctctgc ctcctgggtt caagtgattc 53760
ccctgcctca gcctcctgag tagctgggac tacaggcgcg tgccaccaca tccagctaat 53820
ttttttattt ttagtagaga ctgggtttca tcatgttggc caggatggtc tcaatctctt 53880
gacctggtga tctgcctgcc tcggccccca aaagtgctgg aattacaggc atgagccacc 53940
acgcccagcc gatcttgttc tttcttatga ctgtgtagta ttccatggta tatatgtacc 54000
acattttctt tatgcattct atcattgatg ggcatctagg ttgattccat gtctgctatt 54060
gtgaacagtg ctgcagtgaa cattcacgtg catgtgtctt tgtggtagac tgctttatat 54120
tcctctgggt atatgctcag taataggatt gctggattga atggtagttc ttcttttagc 54180
tctttgagga tactgctttc cacaatggtt gaactaactt acactcatac agtatataag 54240
cattcctttt tctctacaac ctcgccagca tctgttactt tttgacttat aggtgggagc 54300
taaatgataa gaacttatga atgtaaagaa ggaaacagac actgggatct acttgagtgg 54360
gaagagtggg aagagggaga ggagcaaaaa agataactct tggttactaa gcttaatact 54420
tggacaatgt aataatatgt acaacaaacc cctgtgacac atgtttatct atgtaacaaa 54480
ccttcacatg taccccctaaa cctaattttt tttaaaagaa acagaatgcc agccaggcat 54540
agaggctaat tgcctctaat cccagcactt tgggagggtg agatgggcag atcacttgag 54600
cccaggagtt taaggccagc ccaggcaaca tagcagaacc ccatctcttc aaaaagtaca 54660
aaaattagct gggcatggtg gtgtgcacct gtagtcccag ccccttggga gactgagctg 54720
ggaggatggc ttgagcccag gaggtcaagg ctgcagtgag ctgtgatcat gccactgcac 54780
tccagcctgg gcgacacagc aagaccctgt ctccaaaaaa aaaaagaaaa ggaacagaat 54840
gttaccagcc caactgcact tcttcactcc cccaccactc taatgaagtc atcactaacc 54900
cacttctaaa gtactcacat accctatgtc tatggaggta tgtcagtgga ggctaaggta 54960
tgccagtgga ggcttatgta ccttatgtct aaatatttaa agttattaaa ttaaaaaacc 55020
attaaaatat gctttctacc ttgacaaacc tttataacaa aattagaaaa tgtttaatgt 55080
tatggcatta aataattgaa agcaaaatat caaagatgat agaatttaat taattatttt 55140
attttatttt atttgagaga gggtctttct gtgtcaccaa ggctggagtg cagtgatgca 55200
atcatggttc actgcaacct caacttcccg ggctccagtg atcctcccgc ctcagcctcc 55260
caagtggttg ggactacaga catgtgccac caaatccagc taatttttaa attgttttta 55320
atagaggtaa gggtctcact atgttgccta ggccagtctc gaattccagg ggctcaaggg 55380
atccttttgc cttgtcctcc cagagtgctc ggatttaagt tgggagccac tatacccacc 55440
caacataatt caattattta atatttcaca tgttttagta ttcctttgat agggatgtga 55500
tgtttgggtg aataataaag taaatcaaag acatatattt gaaaattatg tagttattct 55560
aaaaaattaa ttatttacct ttattttagc aaaatcagtg tgttagcata atcaagatat 55620
tttggtattc tagtaacaag atctagtcac agtaatgatg taaagattaa aaaataaaat 55680
ataataggaa ccagtataaa caagtggaat ttaaatttaa aatgcaatac cagctgggtg 55740
cgatgcctca cgcctgtaat cccagcactt tgggaggcca aggcaggcgg atcacctggg 55800
gtcaggagtt caagaccagc ctgaccaaca tggaaaaacc ccatctctcc caaaaataca 55860
aaataagttg ggtgtggtgg tgcatgcctg taatcccagc tactcgggag gcggaggcag 55920
gagaatcact tgaaccaggg aggcagaggt tgtgatgagc cgagatcaca ccattgcact 55980
ccagtttggg caacaagggc aaaactctgt ctcaaaaaca aaaagaaaca aaaaacacag 56040
taccatttac attagcaccc ctcaaaatga aatacttagg tataaatcca gcaaaatagg 56100
tataagagat atataagtaa aactataaag ctctgatgaa agaaataaaa gaaccaaata 56160
aatggacaga tattccatgt tcatggatag gaagactcag taatgtcaag atgtcagtcc 56220
ttttcatctt tatctataga ttcagtacac ttccaatcaa aatccccgct agttattttg 56280
tggatattga caaactgatt ctaaagttta tgtggacagg caaaagaccc aaaataggtg 56340
acatgctatt gaagcagaat aaagttaaag tctctacttc ttcttcaata tcattttctt 56400
caatatcaaa tcaactcatt ttcttcaata tcaagtcaac tgtagactga cagtgtccta 56460
cttcaagact tactgtaaag ctacagtaat cagaacagta tggtattggt gaaagaatag 56520
acaaacagat cgatggaaca gaatagagag cccagaaata gacccagaca aatacagtta 56580
atctttgaca aaggagcaaa ggcaatagaa tggaggaaag atacccttt caacaaatgc 56640
tgttggaata actggatgtc cacatgtaaa aaaatgaatc tagacacaga ccttacaccc 56700
ttcacaaaaa tgaacacaaa atggatcatc aacctagaca tgaaacacaa aactatacaa 56760
atttctgctc tgtgaaagac aatgtcaaga gaatgagaag acttggaaaa tatttgcaaa 56820
agacacatct gataaatgat tcttatctaa aatatatagg aactcttaaa actgaacaat 56880
aaaaaagaaa cctgatttta aaatgtgcca aacactctaa tagatatctc atcaaagaag 56940
atatatagat ggcaaataag catatgcaaa gatgctccac atcataggtc atcagaaaaa 57000
tgcaaattaa aatgatgaga ttctactaca cacttattag aatggccaaa atccaaaaca 57060
ctgacaacac caagtgctgg tgaggatgtg ggacaacagg aactttcatt tattgctggt 57120
gggaatgcaa aatggtatag ccactttgga agacagtttg cagtttctta caaaattaaa 57180

cacactcatc atatgatcca acaattgcag tccttgatat ttacccaaag gagttgaaaa 57240
cttatattca cacagaaacc tgcacatagt tgcttacagc agctttattc ataattgcca 57300
aaactttgaa gcaaccaaga tgtccttcaa taggtaaatg aataaataaa ctgtagcata 57360
tccagaaatg gagtactcag tgctaaaaag gaatgaggta tcaagctatg aaaagacatg 57420
gaggaacctt tatattttta ttttttttaga gacagggtct tgctctgtca tccaggctgg 57480
aatgcagtgg cacaattatg gctcactgtt gccttgacct cctgggcttg agctctcttc 57540
ctgcctcagc ctcccaagta gctgggacta caggtgcatg tcaccacacc tggctaattt 57600
ttttttttga gagattgggt cttgctgtgt tgcccaggct ggttttgaac tcctgggctc 57660
aagtgatctt cctgcctcag cctcccgaag tgctgggatt ataagtgtga gccactatgc 57720
ctgacttttt tttaaattta tttttcttct agagacaagg ccttgctttt tattgcccag 57780
gctggagtgt agtgatgcag tcatagctca ctgtggcctc aagatcctgg gctcaagtag 57840
gagcccagct aatttttttt aaaatttttg tagagatgga gccttgcaat gtggcccatg 57900
ctatggagga aacttaaatg catatttcta agtgaagaag ccagtctgaa aatgttatat 57960
accatgggat ttcaactata agacattctg gaaatggcaa aacgaaggca acaataaaaa 58020
gatgaattgt cagggagttg gtcaggggag aatgaatagg tggagtacag aagattttta 58080
gggcagtgaa atgtctctga tacagtaatg gtggacacac gttgtatatt tgtccaaatc 58140
catagattct ataataccaa cagtgaacct taatataaac tatggacttt ggggctgggc 58200
atggtgactc atgtctgtaa tcccaacact ttgggaggtc aagatgggag gatcacttga 58260
ggccaggagt ttgaaaccag cctggtcagc attgtgaggc cctatctcta caaaaataag 58320
gaaaccatgg actttggatt ataatgatgt ttcagtgtag gttcctcagt tataacaaat 58380
gtactactct ggtgggggat gtttataata aaggaggcaa tgcatgtgtt ggggcaggag 58440
gtatacgaga aatctgttta ccttcctcta aattttactg tgaatccaaa actgctctta 58500
aaaaaaaagg tcttaaaaaa taaaattata tttgagggaa aatatttgaa ttattattat 58560
tattttcttt ttgagacgga gtctctctct gtcccccagg ctggagtgca gtggcacgat 58620
ttcggctcac tgcaagctct gcctcccggg ttcacgccat tctcctgcct cagcctccct 58680
agtagctggg actacaggcg cacgctgcta tgcccggcta atttttttgta tttttagtag 58740
agatggggtt tcaccgtgtt agccaggatg gtctcgatct cctgacctcg tgatctgccc 58800
gcctcagcct cccaaagtgc tgggattaca ggcatgagcc accgcgcctg gcctatttga 58860
atgattttta tcaaagatgt aaattaaaac aaatgtaaaa ataaaaaaca aatcactgtc 58920
tgattctatt tgtataaatg tctagaaaat gcaaactaac ttataatggc aaactactct 58980
atagatcagc agttgcctgg aggcaagagg gaagaattgc agtgaggtat gataaaactt 59040
ttggggtaat aaatatattc attatcttga ttgcagtgat gttggcacag gtacatccat 59100
atgtcaagat ttcttgttga atactttatg tagtttattg cataacaatt ctataaaatt 59160
aaaaatcata aaatttttgtt tgttttaaaa acattctttc tttctttttt cctgagacgg 59220
agtctccctg tatcacctag actggagtgc tgattgcagc cttgacctcc taggctcaag 59280
tgatcctcca gcctcagctt cctaagtagc tggtaccaca caggcgcatg ccaccacacc 59340
catataattt ttaaattagt ttttgtagaa atagggtctt gccatgttgc ccaggctggt 59400
cttgaactcc tgggctcaag caatcctcca tcttggcctc ccaaagtgct gaaattacag 59460
gtgtgagcca ctgcacctgg ccatcttaat ttttaatatt taaaagaaaa gtaagggcca 59520
gacactgtgg ctctcacctg taatcccagc ataagaccag cctgggtaac atggcaagac 59580
cccatctcta tcaaaaattg aaaaattaac tgggcatggt ggtggcctgt ggtcccagct 59640
actcaggagg ctgtagctgg aggatcactt gagcctagca cgttgaggct gcaataagcc 59700
atgtttgcat cactgcactc cagcttagat gagagagtga gaccctgtct caaaataaat 59760
agatagataa tatatgtcgt agttttaaaa atatattatt aagataaaaa gcaagccaag 59820
acaattaagt gggggaagaa tagtttttcc aacaaatggt gctggaacaa ctgcatagcc 59880
acaggaaaaa gaatgaagtt agatccccta cctcacacca tataaaaaaa taactcaaaa 59940
tggattaaag acctaagtat aagctgagac aagaagatta cttgaggctt ggagttcaag 60000
accagcctgg gcaacatatt gagacctcgt ctcttaaaaa agaaaaaaaa tcagccgggc 60060
atggcagtgt gtacctgtat tcctagctac tcagaaggct gaggccagag gattgcttga 60120
gcccaggatt tagaggctgc agtgagctat gattgcacca ctgtactcca gcctgggtga 60180
cagagtgaga ccttgtctgc tccacccctc ctccacaaag tgtaaaggta taaatgtttg 60240
tggccttgga ttaggcaatg gtttattaaa tatgacatta aaagcacaaa caacaacaaa 60300
atagattaat tggacttcat caaaattaaa acctctgtgc ttcaaagggc acaccaagaa 60360
agtgaaaaga gaatccacac aatgggagat aatttttttgc aaatcatgta ttttacaaga 60420
ctggtgtcca gaatatataa agaacacttg caactcagca ataaaaagac aagtaacaca 60480
atttaaaaat gttgaaagga tttgaataga catttcttca aagaagacat ataaatcacc 60540
aatgagcata tgaaaatgta ctcaacctca ttggtcatta gagaaatgca aatagaagtc 60600
acacccatta ggatggctaa aataaaaaaa gatgaacaat aacaaatgtt ggcaagtatg 60660
tggaaaaatt agaaccctca tacactgtgg atgggaatgt aaaatggtgc agacactttg 60720
gaaagttggc tattcctcag agatttacca catggcacag caattctact tttaggtgta 60780

tacccaagac aattaaaaag atatatacag gccgggcgcg gtggctcaag cctgtaatcc 60840
cagcactttg gccaaggtgg gtggatcacg aggtcaggag atcgagacca tcctggctaa 60900
tacagtgaaa ccccatctct actaaaaata caaaaaatta gctgggcgtg gtgggggggc 60960
gcctgtagtc ccagctactc gggaggctga ggcaggagaa tgtcgtgaac ccggcaggcg 61020
gggcttgcag tgagccgaga ttgcgccact gcactccagc ctgggcaaca gagcgagact 61080
ccgtctcaaa aaaaaaaaag atatatacac agaaaaactt gtacatagat gttcatagta 61140
gtattccaat aaacatgccc atcagtagat gaatggataa gcaaaacgtg gtgtattaat 61200
aaatgaaata ttatccagcc acaaaaagca atgaagtact gatacatgat ccaatatgga 61260
tggaccttga aaactatact aaatgaaaga aaccagccac aaaaggccac atagtacatg 61320
attacatttg tataaaatgt ccagaattag caattccata aagacagaaa gtagattagt 61380
agttgccaag ggctgaggga aggaggaatg ggagtgactg ctaatgggta cagggtttct 61440
ttttggggtg agaaaagtgt tctggaatta cataatgatg atagttgtac aactttgtga 61500
atatactaag acacactgaa ttgtatcttt taaaaagtta aattttatgg tatgtgaatg 61560
atacctcatt aaaatagtta catgagaaaa aaatcaaagg caaaatacag agtataattc 61620
aagtatttta atttttaaat ataaagtatt tatagccaaa tttgatttac ttctaaaatg 61680
tcttattaaa tagtttaata aaagcaaaac tgttctagca ttcagtgttt attaatttgc 61740
aatacacaaa caatatcatg ttttactcat gttgggtcca ccacatatgt atatatttaa 61800
gaataatgtg attggtcagt actgcaaaat gttttcgtgt acggtggctg tgagtaccat 61860
actaattagt acactaagaa ctatgaattg gaacaggaag aaaagcaaga aaatgaacat 61920
tcagcactat ttggaaatga aacctcacta tgcaagaatt cattgcattc atgccttctc 61980
agggggagtg tttgataaat taatatttca ctaacctaag agcttctgct gctcaaatcc 62040
tctacacaca ctaatgcagt gtcagtctct gagtttggca gtggcacagc cataacttta 62100
tggcagttag gtgaaatccc tttttgtttt gtggacatag aacaacagat gtggagttgt 62160
acttctctgg ggccagcaca aacgccatta agtctgtgag tctcattctg tcagtgctga 62220
gcatccatgt gttctttgat ttgcgagtgt gtgtttgtga ttggcaggtt tctcttacgt 62280
gactctgaag gagtgttgct gatcacatca cacccatcct taggccttat tgatcatcag 62340
tgtaccttcc cactactata ctttaaatag atgcctgtta tttaaatttg attttgagat 62400
taaacagaat ggcagagaca attttgagac acatctttct ttcatgctct ggtaggaaga 62460
tcaagatttc taggacagta gaacagagta aaagatgagt gctgttggga tcctatcttt 62520
ctcctaagct atttttttctc tctcagttat tcatcatcta tctcaatttt tcctaatgaa 62580
ctcttctata taaaagagga tccaggtccc acatccactg tcaaggagga atgtaagatt 62640
gacttgcaac tcagcctgtg tacgcagttt atggtttttc ttggcaggtt taatgttctt 62700
tcttatctct taacttcttg ctattcaata gtaagtaact ccctggcaga attacctgtg 62760
gctagagaat gctgttatca gagcatcttt gtttaattgg tacttagaac agaaggtgtc 62820
acctatttga caggccaaca attatgagca aggaggcatt tgattcatca agatagaaat 62880
ctgcctgtta ggtggaaaca tgtctatgtg ggttgatatg tttttagaat attaaggctt 62940
gtttgtgcat gacaacttta ggaaggtgta ctccaaatgt ctccaaaggt ttgctgtagt 63000
tctttacaga agttgggctg ctcctggtgg acagtgtgta acagtgaaca atgtatgctc 63060
tagactgggt tcccttcctc caccctgtgt ctgtgtggcc ttgggcaagt tgtttaacca 63120
accacttttt gcctcagttt ctttatcgga acaaggagaa taagaatact tcaatcaggc 63180
caggcgtggt gactcacgcc tgtaatccca gcctttggga ggccgaggtg ggtggatcac 63240
ctgaggtcag gagttccaga ccagcctggc caacatggtg aaaccccatc tctcctttac 63300
ttatgctggc ctgatattga tcgtccatgg tagaattgat actgcttgac aaagcagcct 63360
atttcagtca ggacccttct tctctagttt cctctgtagc tattacctta gccttccatt 63420
tcattcttca cactacagat actctcattg ataaaggaat gatgtcttta tgctttcaag 63480
cattctggca agttagtaat tcaactatga ttctaggtca gacaaaacca gttatgaaca 63540
taagactgtt tttaatctcc tccctggtcc cccaaccacc caccccaatc aggagaaact 63600
atgttctgca ttggtttaag gaacccgctt ctttctttga tacctgacct acagatccaa 63660
tctattccca ggaattttga taagaattct caaatcctca gcaaggctat gccactgtca 63720
tgactctcct attcctggta gtgtcattct cagtgtaggc tgtttgatag gtagttttgt 63780
gaagtcttgt tcatcataat ggatcatatg atttttaaaa gcaggacctg ggtcatatgc 63840
ccagattaat ttcaacaaag ttggtatgtt ttcttcctaa aattaatttt ttatgattat 63900
caaagtttta tatgcataca gtataaaata tcagatttct ataaggctta taataaagag 63960
tagcattcct ctgccctatt cccctctcat actgggcacc ttactccaaa atcaactgct 64020
ttcagctctt ttagccattt cttatggtta tcttcatatt tcaaaacagc atgcttatat 64080
ggttgtattt tgaattttca aacttagatt tttatctaca gacttcctta tggaagataa 64140
atatttaact ctttttgtgc tactcctttt ctcatctctt aatgtaggct atgttaaaat 64200
ttttggttaa atcaattaaa agccagtata gtatagtggc ttaagagtga gggtgctcac 64260
ctcaaattcc agcctcacca cccattactt tgtgtgactt tgagcaagct ttaaaacgtc 64320
agtgtctcag ttttgtcaac tgagtagata cctcatagaa ttgctgttga tattaagtga 64380

cttaatccta tgggctgaat ttgtctccca aagttcattt gttggaaact taatccccag 64440
tgcatgtgtt aagaagtggg acctttcaga gttgaatagg ctatgagggc tctgccttca 64500
tgaatggatt aatgccgttg ttgcagtagt gggttcctta taaaaggaag tgtttgaccc 64560
cctttccttg cctcctcatg catgtgatgg ccttagccat gttataatgc agtagtaacg 64620
cccttaccag acactggctc cttgatcttg gacttctcag cctccagaac tgtaagaaat 64680
aaaagttttt tctttataaa ttacccagtc tctggtattc tgttatggca gcaaaaaaca 64740
gactgagaca cttaatatat atgaagcatc tagactgtct ggcacattgt acattttaaa 64800
tcccagatat cgatatcatc aatatcatca tcatcatcat ctgtggctgt ataatacctc 64860
cctctgcatt taaaggatga gggctggtgt agcagttatt aatataagtg agctagttgg 64920
ctataaacct ctcctatagg ctttgccata aacttgtgtc atggataatc agagaatttg 64980
gacctcctaa tgacaggcca ctgacaagaa aagcccagag ggagctgatt gagcatgctc 65040
agttctttct accaaaggcc tcaatcagac agattctctt tcccaggagt agacactgag 65100
ggggtggaaa caggctctga gttgactcca ctgagaagtc cctgaatgga tagcagccag 65160
ggaattaagg agtttccctg tggcaagtct ctactcgtac aatattagga cagctgtttt 65220
tttaatttgt tctgtgggtg catatttttg atctccacaa cataactact tactcactgt 65280
gttgcttttt gaaatttata acatctaaat ttgcaaatgt gaggtatgct ctcaggaaaa 65340
ttatatatct gttaaaattc tttgctcctt ttaggacact tccatccatc tggtgacctc 65400
tagtagcgta catactcagc aacaattaaa ggtagtttct ggcttgtttt tcccaaacaa 65460
taatttgttc aaaatacaga aatttggaaa gtaccccatg atgagaaact tttgaaagag 65520
aatttaatgt acagtaacta tttttttcct ctgaggaata attttgggaa aagaaatttg 65580
ctttatattc aggcatattg agtaagttgc cttcattttc aaacaaattc atcacctctc 65640
ctttccccat tccccaaact tgttcatttt cttcattctt aaactctgca tctgtcttcc 65700
cctcacttta ccacgtagct aaatctctta agagccaacc tattgccact cccttcatct 65760
tttttaatct aatggatcat aagtcctgtt gatttcatca ctcacatctc tcaaatcggt 65820
ctccttggat atcttgctac ttctgatttc agacaattgt catttgtcac ttgcagcagc 65880
cttccagctg ccttcctccc aatctccctt atgagcattc tgttcttcaa actgttgcca 65940
aagaggaatt tctaaaatac cgtttgatca tgtaattcca ggttttaaaa acctaatgtt 66000
gactctctga catttaaaga ataaagtccg tatttcttgg gatgactgac atggctttgc 66060
tgtgattctg gctacatcct tagagaatca cttgcagata ttaattttgc atatgctatt 66120
ccctttacat ggaatgttcc tgcctcctct ttcccccatc ctttgcctga tgaacgccta 66180
ttcatcctcc atgtcacttt gggtgttacc ctgaggaaac agtggttatt ttttccctgc 66240
ccatctgggc tggatacccc tcctaagttg tcccacaaca cacagtacac accttggtcg 66300
attatcacac tgtttgttgt tctgccgttg tttgtgtctc ttgaatatga gtttcttgag 66360
attagggacc tgagatcccc agtgcccagc agagcaggac ctgtttccac cccctcagta 66420
actactcctg gggaagagaa cctacaaata aatgtgtatt gcatgaataa agctatatat 66480
ccccttcctg ttcatactta tccatttaat ttctgaactc taaatgctgt atttttctcc 66540
attaaatttg gttttattgg tttcagtttc ctctttatat ggattttgag tcctaatttt 66600
gtcaaccagc atatcaacag ttctttctag cttttgtccc ccggtagaat ggcagctcca 66660
tgaagacagg gattagtgtc tgttttgttc actgctgttt tctcagcatc tagaatagtg 66720
ctggcacata tagatactca caaagtattt gttgaatgaa tgaaggtgtt gcatacaaat 66780
ttgataaata aaacttaggt tttctttcaa atattaatcc ttagctccat tctgttcaat 66840
atttttatta gagataacct ttaaaaatct tccttgtcac ataagagtta attccatgaa 66900
tcttacagaa caaggctata ctgagaaatt caaggaacac cttaatgaat cagggttatt 66960
tcattgtgag agagtataga aatggttgat agtatgctgt aagtaatttt actttacgtg 67020
taagtacttt tctctggctt tccagaacat gctgttggag taagagagga atgccttatt 67080
gtggaccgag gggatagatt tggatacagc cttctttgag gaaggtagag aggtcaacta 67140
ttacatatcc agcagtaact tccctttcaa agactaagtg ttctcattca atcatttgat 67200
acttttgtg catctactac aatttgaaga atagaagagg gaaatgcatg tgtaaggcat 67260
ggtggcaaca tttaagaaac ccagatttgg ggatacaagg tgtgtgtgtg cacatgcatg 67320
tgtgtgcaat ttaaatgcac agggtaaaga actttgagtc agtgacaatg agtggcaatg 67380
gtggtaatta agtgctgggg aagatcaagg gagagagagg gtgctgagag ccaatagggt 67440
ggaacatgtt tgaaagagct gggttctgaa gtattcctcc atagaagggc attttaaata 67500
gctttttgc gtccttattc tgttaagcat tattaaattg ttctcccatc tttaaaaggt 67560
tccctagctt aggtgcactg gcaagaaatt ataaaagcag catgaccagg cgtggtggct 67620
catgcctgta atcgcagcac tttgggaggc caaggcggga ggatttcttg agttcaggag 67680
tttgagacca gcctgggcaa catggtgaaa ccctgcctct acaaaaaata caaaaattag 67740
ctgggagcag tggcacgtgc ctctagcccc agctactcag gaggctaagg tgggaggttg 67800
gcttgagcct gggaggtgga ggttgcagta aactgagatt gtgccactac acacagcctg 67860
ggcaacagag cgagaccctg tctcaaaaaa aaaaaattaa attaaaacaa taaaagcagt 67920
gtggttgctt ataaggagtg atggagtgga cagaggaggt ttttgggcta gtcaggtaag 67980

gctggggtat gaatctagaa gtttccattt aaatagcagg gagcccttag gcaaatcact 68040
taacttctga gctttgcctg tttcacctga ggttgcttca aagattgaat gaaaccgtat 68100
atataaagtg ctcctaaaca tcatctgcca tgtggcaggt tctcaagaaa tgttagtttc 68160
cctttctcct tacaaagata agatgcttgc tttgagtata tttttaggct tcctgatcat 68220
tattggttat gattttaaat cttggtccag ccaccaccca tatggtttct gcagttaaca 68280
aaagaggcaa aggttcacta ctgggggaaa agaggtgtac agaaatgggt gtagagaaag 68340
tagatgtttg aaggggatct aattaggaaa gtatttttcc tggtggtcca gaatttaaaa 68400
ttatagagtc ttatgagaaa atgataatat tcaggttaag aacagtttat tatttcctct 68460
ctatattgga gaatattttc attatcttat ataagaacct tgtaaaaatt tttcttctta 68520
actagcttcc cactataggc cattaatcct gttattattt caaagattaa acaactatag 68580
taataacagg attatatgtg cattaattta tattacttca ttccgcaaag gatttgaggt 68640
agtttttaga agcataaaac actgcacaaa taaaatagta ggataggagt agaaaataaa 68700
atttcagcaa ccatgaaaaa tatgacatag tatatgttgt taagactggg ggaatgtaaa 68760
cacatcacca gtcagggcct atatagttgt tacagtccca tagcaaattt gactctaagc 68820
ttctggcaaa caacatgaaa agggaagcat gatggatgtg gttaaataag acacaattta 68880
cttgttactt tactcaagga agcaagtatt ttttgccctt ttgtttctta taggagatgc 68940
tgggtaatga agtaatgttg gcattggcct tatagtggag gtaataatgg attttatcag 69000
gcagtttctt taagcatctc ttgatgaaag atgaggctat gacatcaaga gacaattctg 69060
aggccgggag cagtggctca cacctataat cccagcactt tgggaagctg aggcgggcag 69120
atcacttgag gtcaggagtt cgagaccagc ctggccaaca tggaaacctc gtctctactg 69180
aaaatacaaa aatcagaaac cctgtcttta ctaaaaatac aaaaattagc tgggcgtggt 69240
ggcagtgcct gcaatctcag ctacttggga ggctgaggca ggagaattgc ttgaacccag 69300
taggtggagg ttgcagtgaa ctgaaatcac accactgcac tccagcttgg gtgacagaac 69360
aagactccat ctcaaaaaaa aaaaaaaaaa aaaaagacaa ttctgaaagt ggtgagtcgt 69420
tctaccaggg gccaggatga tctcatctgg gttatggatt ctaggtctgg cctcatctaa 69480
ggtgacacac agagggtaca ccgcacaatc ttcctatctt ctttaaatat cactgctttc 69540
aacaggactt ttttttttt tttttgaaac aggatcttgc cctgccctgt cacccaggtt 69600
ggagggcagt ggcatgatca tggctcactg cagccttgac ctcccagctc agccttctga 69660
gtagctggga caacaggtgc atgccaccgt gactggctaa ttttaaaaag ttgtttcgtt 69720
ttttttttc tgtagagaca gggtctccct agattgtcca ggctggtctc aaactcctgg 69780
cctcaagcag ctctccagtc tttgcctctc aaagtgttgg gattacagac gtgagccacc 69840
acgcctggcc caaaaaacat aataatgtgg ttattcgaag aagtgatttc ctctcaaaac 69900
ataaattcat ttcttctttt actcttgtaa ctttctgagg tgaaaagtag gaagttccca 69960
gatttttcat tgctgtaaag aaagattata gacaagcaag gaaggagtta gaataacctg 70020
tgtgataatg aattagaaga gttggaggtg tatgtaagta tgctcagcat gaatttatgt 70080
ttagcttaat gtagatacag atggttacat gtggaaagta tttatagcta tgcatagata 70140
ggttggtata tgtacatgta ttttctacct cctttggcta agagggcgca gaagccatga 70200
catccctgtt gcaacaaaaa cacctagcac catttatctt ggtttgtaat actattctcc 70260
agtaaaaaca accagggctc cttgcagaaa gggctgatga taatatataa gattagcctg 70320
gagcatctta tatatcagaa agcaagggag tactcaaaaa ctaaaaacaa taaactgctc 70380
cccaataatg ggagtatgtc aaagggtcac aggagccatg tgaaagagtt tgcaatagcc 70440
aacaaaatga agaagtattt gaatttaaat cagagtataa aataaatatc tacgagtcca 70500
taatgatata aacaagtgat tgaataaata aataaatgtg ggagactaga caaatccccc 70560
atgcagaaga attccaagta atttatgtag gtaaatactt cactgtcaaa gaggcagagc 70620
ataattaatt ccccactcct gtaagtgtgg gctgcttagt gacttccttc caggagtgca 70680
gtataaatag ggaaaagaag acaacttcac agtggaggaa tctggcaaac tctgtcccac 70740
ccagatgatc aaggtttaca tcaaaagcac taagccatgt tgatactgca ctgtgtttcc 70800
ttgatataat gggatgaaat ggcactttgt gtagtcgtgc ttccaaaaaa cctcatagcc 70860
ctagactaat catgggaaga acaccagaca aatcctaatt gagggacatg ctacttaatt 70920
cctgaaaagt actcctcaat gctgtcaagg tcatctgaaa caagaaaagt ctggcaaact 70980
gtcacagtca agaggagact aaggagatat gacgactaaa tataatgtgg tatcttgtat 71040
gggatcttgg aatagaaaaa ggatattagt taaaactgag gaaatctgag taaaatatag 71100
atgtttgtta ctaataatgt atcaatagta gttcattaat tgtgacaaat gtaacatact 71160
atcgtaagat gttaataata gaggaaactg gatgtgaggt atatgggcac tccctgtact 71220
gttcacaatt gttatgtaaa tctgaaacta ttctaaaatg aaagtgtatt ttattttatt 71280
ttattttgag acggagtctt gttctgttgc ccaggctgta gtgcagtggt gaaatctcgg 71340
ctcactgcaa cctccacctc ctggtctcaa gcgattctcc tgcctcagcc tcctaagtag 71400
ctgggattac aggcacgcgc caccacaccc agctaatttt tatattttta gtagagatgg 71460
ggtttcacca tgttggccaa gctggtcttg aactcctgac ctcaggtgat ccacctgcct 71520
cggcctccca aagtgctgag attacaggcg tgagccaccg cgcccagaca aaagtttatt 71580

ttttaaaagg taggaaagtt tcacattttg atcgtactat tgagataaaa cttggttgtt 71640
gttgttgttt ttgagatgga gtcttgtact gtcacctgga ctggagtgca atggcatgat 71700
cttggctcac tgctacctcc acttcccggg ttcaagcgat tctcctgcct cagcctcctg 71760
agtagctggg actacaggca cctgccacta cgcccagcta atttttttgta tttttagtag 71820
agatgggttt cattatgttg gccagactgg tctctaactc ctgacctcgt gatccgccta 71880
ccttggcctc ccgaagtgct gggattacag gcatgagcca ccgtgcctgg ccaaacttgg 71940
tgttttcatc tgagaatgag gtttccaaga attgaagata caagttagct aaatagtatc 72000
agtgaaacca cagagtataa cttcagacca cttgtggttt taaggcagat ctatgccaca 72060
gagaagtatt tgaatttaaa tcaaagtata aaataaatat ctatgagtcc ataatgatat 72120
aaacaagtga ttgaataaat acataaatgt gggagactag acaaatctcc catgcagaag 72180
aattccaaat agtttatgta ggtaaatatt tcagagtatc agtactatga tccttcttac 72240
aatgctatat ccgactaaaa aaaaaaaaaa gaaaaaaaga aaaagaactg tgatcctagt 72300
tgcgaatgtt gacttatagc cagaaaagat gactttctta tttgaaagca tcatgaaatg 72360
ctatcagcat gcagaaatga gagggcataa cagagagcta ttctatcgta tttttattta 72420
atgttttaaa agagctatta ataatataat gtgaagagaa gttttaattt cagaggaatg 72480
ggatgtttgg ttttaagatt atttttgcaa aatgtagtct gtcatttta aagtgaactt 72540
tataatgaaa agagtttgaa attatttccc tattattaat atttatttat ttttattact 72600
tttctgccat gacaattaca gtgcacattt tccctatttt tagctttttt ctggtaaagt 72660
gattttacct cctgtagatc gtcagtcccc agactttctg gcactaggga ctggtttcat 72720
ggaggacaat tttttgacaa attggagggg gaatgatttc aggataaaac tgttctatct 72780
cagatcatca ggcattagat tcttacaagg aatgtgcaac ctagatcccc tgcatgcaca 72840
gttcacaata gggtttgtgc tcctatgaga atctaatgct gtgctgatcc gacaggaggt 72900
ggagctcagg tggtaatgct cgctcaccta ccactcacct cctgctgtgt ggcctggttc 72960
ctaacaggct cgggggctgg ggacccctgc tgtaaataac ctttcgaaga tctgaaaatt 73020
aactttagta ttttttgta tttactcgat attttaaaca aacaaaaatc tagagaatga 73080
cataacaaat gtattttctg cgtatccact actctttact cagatcctaa cattctcagt 73140
tttgaaaaga tgtaggctgg tcacagtggc tcatgcctgt aatgtcaaca ctttaggagg 73200
ccaaggagga cagatcacgt gaatccagga gttagagacc agcctgagct ataaggcaaa 73260
accctgtctc tacaaaaatt agccaggtgt ggtggcctgc acctgtagtt ctagctactc 73320
gggagcctga ggtgggaggg tcaattgacc agaggctata gtgagctgtg atcctgccac 73380
tacactccag tctaggtggc aaagcaagac cctgactcaa aaaaaaaaaa aaaaaaaaaa 73440
aaaaaaaaag atgtaaacat tacaggccca gctgtggtcc cgtgtacact tctccattcc 73500
taatcccctt catgcacact ttcccacagg taatcactat cccacatttg ctttaattat 73560
tcccatcatg ttttatgcta tgactacaaa tgtgtgaatc tatattccac tgtatgttaa 73620
taccaacatt ttttcattta ttttaccgct gatgggcaat ttaggttatt tccagatttt 73680
tactattaca gatgcttctg tgaacatcct tactcatttc tccttgtgca catgtgcaaa 73740
atttctctgg agtacatacc tgggagtgga attactagaa tgtggtgcat gtacatgttg 73800
ttgtatttat cagtagttca ctcttattta ttgctgtgta ttccattaca tagatatacc 73860
ataatttgtt tatctaatca cttatttata gacatttggg ttctttctgt ttttttgact 73920
attagaaata aagcagctat aaacatttgt atataagttt ttgtataggc atatgctttg 73980
gtttctctta ggctcagggg ctgggaaccc cctaggaatg gaggaatagc tgggtcatat 74040
ggtaggttta tgtttaactt tttaagaaat tcctgaactg ctttccaaag tgtttgtact 74100
attttatatt cccagtagca gtatgtgaaa gttcttgttg ctctacatcc tcaccaacac 74160
ttggtatgat tagtcatttt tagatatcct gtgatgtgtg tggtttttca ttgttttaat 74220
tgacagttcc ctgatgacta atgatattgg gtgtcatagg catattttcc atttatatat 74280
ctttttgta aagttcaaat tttttgccca tttttaaaat tgagttgctt ggtttttatt 74340
gagttctgag aattttgtat gtaatttata tacaagttct tactagatat gtaatttgca 74400
aatcttctct ctcagtttgt ggcttgtctt tttattctct tagcagtgtc tttcaaagag 74460
aagttcttag ttttgatgaa gtttagttta tcaacatttt cttgtactga tttcttgtat 74520
gggcatttag gtttttcatg ttatatctaa gaaatctttg cctaaaccaa gatcacaaag 74580
attttctccg gtatttttctt cttttgtttg agacagagtc tcactctgtt gcccaggctg 74640
gagtgcagta gtgcagtctc agctcactgc aacctccacc tcctgggttc aagcgattct 74700
tgtgcttcag cctcctgaat agctggtatt acaggcatgt gccaccatgc tcagctaatt 74760
tttgtatttt tttaagtaga gacagggttt caccatgttg ggaaggctgg tcttcaattc 74820
ctaacctcag gtgatctgcc caactcagcc tcccaaagtg ctggtgacac aggattttgc 74880
tcagctactt tgccaaccag ggactcactc ggccatgggc caaggcaccc cgctcacttg 74940
gtccacctgt gctatagctt ctactgacgt tcagcggttc ccgagctctt gtcatgcatc 75000
taaaaagaag gaggatatgc tgacaatttg aagtgtaagg atgggtggag aagaatttta 75060
ctgagttatg gaacaactct cagcattaag gggacacggg gtgctccctc acccccacag 75120
tcaggtggtt tttctctctc tctctctgtg tctgggtctg gggcttttta tggactcaga 75180

atggggagtg tgtacagatt ggtttgtgag tatgcaataa aagttaaagt gaagacaata 75240
ttcaaaggtg ggcgcgatag tgtagaaaac caattaggaa agggtaggta tgtagcctgg 75300
catggtggct cgtgcctgta atcccagcac tttgggatgc caaggcaggt ggatcacctg 75360
agatcaggag tttgaaacaa gcctgaccaa catggtgaaa ccccgtctct actaaaaata 75420
caaaaattag cagggtgtgg tggcacacac ctgtagtccc aactacttgg tagactgagg 75480
cagaagaatt gctggaacct gggaggcaga ggttgcagtg agccaagatt gcaccatggc 75540
acttctgcct ggctgacata gcgagacagt gtctcaaaaa aaaaaaaaaa aaaaaaaaaa 75600
agaaagggta ggtatgtgta aaataggtgg agggtgggga tcaatcagag gaaagcatgc 75660
caaatgggaa gacaggttct caatccagtc catggatttg cctgggactt gtagctaggc 75720
tttaaagtgt cttcagcttg aaggtcaggt ttcaccaggg acccgttcct acacccctat 75780
ctgcctaggc atttgtctgc ctcttgctgc tatcagttct cccctctgaa gaggtacatc 75840
taactgccat tagaataggg atgatgactg atactaactg cttcctgctg acaggggtgc 75900
ggggtgctgt tttgggaaaa tggcagtcag atctccctca gaggcctatc taagggtccc 75960
cagtaaaagg gagccatcgt ctgaggctcc agtttcatga ctggagttta atggcctgaa 76020
aatgagaaga caaccagatt attagaaggc atgtatcaaa acaaaataag ggggtaagga 76080
cagctcaaaa atcctgaggc tgccaacatg cccagataac aggtggctat agttatgcct 76140
gctaagattt gggtgaatga ggctcggctt tggtcagctt ctttggtctt attttcccaa 76200
acaaagaaac ctctgggtta cgggcaccct gtttactcct atcacctggc aggatttgca 76260
ggataattgc ccagaactag aatattgatc cagattttta catcacccat ccctttgtt 76320
tcttctgagc tgcagctgat gatcactggt tggttcacag aaataagcag ggttagtcta 76380
aaatgcagac aaaaacttaa aaacaactaa tgagactaga atttaatgaa aagtgtatga 76440
taaattttga aacataattt ttctctctcc agtcctcatt tttgttaaaa acaaatcatg 76500
ataggactga gtcatttgca gaataaactt tagtcttata tttggcctgg ttatttgcat 76560
aaagcacagc aagaataatt atttttcaca caggctttta aaattggctt tgatggaact 76620
ctgttccaca aggaatttca gataagacct tttaaagctg agcccagcca tgggtttgta 76680
tcctcaaata cctatgagtt gggtaaattc ctctcttctt gaggtcccaa gataacatgg 76740
ggttcctggg cctattagaa agtgacattc tttattcacc acagattagg aactctgtac 76800
agggactgtg tagaagacaa agtatgaggc cagttttccc aaggggcttt tattggttct 76860
gcaagtcaaa cttgatttct taaaggtaag cacacccttc cagtcaaagc cttggtaaaa 76920
caaccagttt ctccagttgt gtcctgttgc aaaagaaaat ggattcttac tgcactgatg 76980
caaataactg tattgctgca aattaagaat actcacaaat agtttccaag ttctgaggaa 77040
accaggcaaa aagaaataaa tgtgctccaa attttgttca ctggagtata ccttactcaa 77100
ttgttaaaag ctatagatag ctcaacatga aagtttcctt gactctgaaa aacaaaacaa 77160
ggatcagcaa tgttttaagc aaaaaaagat tacttcagct ttctattagt tcagtacatt 77220
ctattaactc ttcttctgct tgatattcat gaacatttca gctcttcatg agtcctgtac 77280
attttccctc tattccaatg tcacaatctc caaagttatc agaaacctgc atttgagagc 77340
acctgtcaaa gtcccatagc tgattataaa ccatcttttg aaaaggatca aaataagaca 77400
attgtctgtg aatgacaaaa tgtctttggg taataacagt caaagccatg attgacaaag 77460
aaatttggtt atttctgagc tttacaataa caacataata attttttttt ttttttttt 77520
ttttgagacg gagtctcgct ctgtcgccca ggctggagtg cagtggcggg atctcggctc 77580
actgcaagct ccgcctcccg ggttcacacc attctcctgc ctcagcctcc caagtagctg 77640
ggactacagg cgcccgccac tacgcccggc taatttttg tatttttagt agagacgggg 77700
tttcaccgtt ttagccggga tggtctcgat ctcctgacct cgtgatccgc ccgcctcggc 77760
ctcccaaagt gctgggatta caggcgtgag ccaccgcgcc cggcaacata ataattttaa 77820
ttacgattga tagcatatac tcagacatta gaattttaga aacctcatag aattttggaa 77880
catatgtatt tttcattaaa atataacctg aagaagatta aacattattt ttattttggc 77940
aatcccacat aactaaacat gtcagttaat cctgtttacc tctcttttgg atgctccagg 78000
agccctctgt agtattcaaa agtaaggggt cagaaaagac aaccttgaaa ctgaagtttg 78060
attttgggaa gcctgttaag tacattagag gtttaaaaca ctttatatta tgaaatacaa 78120
ttccagatta ccataagtca tttatttagc caaaatgatg actcaaaaat ttttaaaaag 78180
gtaaaaacct ttactcatta agagtgaaga cagctttcca aacaaacaat ccatctctgg 78240
tctctcccac acccttttatt ttttgatgaa atctttagat aattctgtcc aatcttaacc 78300
agtttgacca tgaggtgaga ttcttataaa cctttacaaa tttttgttaa agagtagatc 78360
agtgccttaa gaaaaccttg ttcttttatt ctaatgttca atttacagaa aaaccatgta 78420
ataccctttt gaatttagtc aatatgttca cacacagaat ttcttttgca agattaattt 78480
ttacaaacct tccacaacct tcaactttat cttatccaac ttaaaacaat tatttaatcc 78540
tctaaactag gcaaaaattt aaattcccat gccttcttat aatcctttac taaaaacaca 78600
tttactttcc ttacacacct tgatgtaaat ctgttttcag tagtctcaat tacatggtat 78660
aatggtaaac cttagcaatt tttaatttta atgtaaagcc tggtaagtta ttttaattat 78720
gtgctaccaa ttatacctta agttgtagcg actctggtgt gctattggta atatggcctt 78780

acaaaactga aaagcaagca aggtgaacaa ttttcaaaag cccaagaagc aggctgggtg 78840
tggtggctca cacctgtaat cccagcactt tgggagaccc agccaggtgg atcacctgag 78900
gtcaggagtt caagaccagc ctggctaaca tggtgaaaat ccgtctctac taaaaaaaaa 78960
ataaaaaaat tagttggtca ttatggtgta tgcctgtagt cccagctact caggagggtg 79020
aggcaggaga atcacttgaa cctgggaggc agaggcaacc aagatcatgc cactgcactc 79080
catcctggac aacagagtga tactgttgaa aaaacaaaaa agagaaaaaa gccaaagaag 79140
cagtttataa ctttaaagca tttagtaaac ctagtatctg acctgcataa tttagaccac 79200
atgtttacat tttgaagaca tttgtatttt accaataatc tctaaaactt tttatttttc 79260
aaatattaaa atcatttgaa ctaaaaggta ttatagcttt tatttttcct tcagaaaata 79320
tttgatctaa gtgcttattt ttctctaagc caattaatta gagctctttt ttatacaaac 79380
atcacacata ttgcacatat ataactacac agacagagga tccagtagtt gtaagatttt 79440
tcattgtcca atctcctaat tagattactg acctcaggat ggagcccttc aagagcaggg 79500
ctaggaaagc atgcagtttc tagggcctaa taaatagtta tagctggaag acaaaaacag 79560
attttgagag ggatttatct gcttttaatt ctttgggttt catgaggaaa acagaggttt 79620
ttttctaaaa tggggtcagt ggtgcctctt ccatttttc cagggagtcc caggccatca 79680
gaagttatct tagggcctct catgcgtgca ttaagagagg caagacaaaa tggagaaaag 79740
taattcagtt gactgaaaaa gaaaatcttt ttccagtgaa acaagatgca agaagaggaa 79800
aacatagagg ccttttaaat atgcctatag cttggatatc cacttttaat taagctgact 79860
ttttaccata gtgctcttat tttaaaaaat ccttttaaat cccttgttac ccaactttag 79920
cctcaccaag tggccaatat ttctggcttt tgaactttac caaaagtaac ctcacaggtg 79980
aaaccaacaa gccttaagta aggttgtgac ttcactgcct gtgtacaagg tattttcaaa 80040
gagatggtaa gcagttttta caaaatctag aatctttaaa gatagctcag agaaaggaag 80100
atttaagaaa ggaagctaga agttgttcat ggaggggaag agaatcagca aatgataaaa 80160
gtcacacaga tattagccag aaagtactca ttccctaagc caggattgaa cctgggcttc 80220
cattgtaaaa tggcagagac caaaagaaag tccttccacg tggttacaag gtcaagctcc 80280
caaggacata aaacaagatg gagaccttat tcagtttttt ttccttcaga gacctgtagc 80340
aaagtctgta actgaccagt ttgcatggct ggcttgaaca gtgggcttat ggggtcctag 80400
gtctgtgttc tatcctgtct tactccttat gacagaactg tacagaaaga cacacaaagc 80460
atagcagatt ggctacagct taagattagc ctcacaaatc ctttttttcca ttaatcaaaa 80520
ctttacagaa gaataaacag tgatttttat ccttcctttt actggtttgc acagggagaa 80580
agaggccaga agtctgactg gtaagaactt ttactctttt actggcatgt caggcttctg 80640
ggttcccttc ccagttcaat tttaagccaa gcagtttaag gtttggggat actaaacttt 80700
tcacagatat tttccagtat gtttaaatag ttttgtttag cccagattta ataatagtta 80760
tctgtagtag ggtttgcatg accactttat tccaatgcca caaacagaag ttgattagca 80820
aattatatat atatttgcat acgtttattt tacataattt aaggctatta gaccaatatt 80880
ggttattaca gagcatagat accttgaaag aggccaagat ggagtggaaa tgcaggttgg 80940
agctaccaaa gggcaaaaaa ctattcatca acctttaact acaataaatc ttccctttgg 81000
gtattttcat ttgtaaacac tttgcccttg cagtcattta catcggcata attacaacac 81060
ctttgcttta tttggcatgc aggagaatct tcttaaatat taaacctaac cactttgagt 81120
gatttgcatc ctgcttttgt catacattta gagccttggg atttttttctc agaatggagt 81180
aggaaacaaa tagggtctgg atagggaaat tgaaaagctt cctggtattt ttctgttgaa 81240
agatttctta acatggcttc ttggatgtgt ctctctgatg tcaaacatac acacatattc 81300
aaataagagt tatacaagca catcttgcac atttttggca tctatgtctc aagacacagg 81360
acattctatc tggtgttctg atcgaaccac ttttgcatgt tactagactg aaaattattg 81420
gaaggtagag aaatctctta tttgttttta tattcccaac agcctaggat agagcctaga 81480
acattaaaga acactaaaat tttaatagtg taactgaaaa gcaggttagt tggtcactgc 81540
atgtagagtc caattaacaa gagcaagttc tgatacaaag aagtgtattt atttcaaaac 81600
tagcttaggg gaagaggcac aaagcatcct gcctttaaat gtgccacttc acctttggag 81660
caaaaagtgg gcatttttat aaggtagggg aggaaatgag caagggcaag tgtccctctg 81720
ctactgggca agtatctgag ctggcacctt cttgggcaga agtaagttgt aaaagtggcc 81780
aagtgggtat gctttcaaca tgccctccta gtgggcatga gttctgagat gaccctgtgg 81840
agagttctgt gggggcatgc tttggtctgc aaatagactg ttaactttcg aggagagatc 81900
cttgggggga aaatatatat taggaagtcc tctgtgggtg ttttgtagaa ggacctagag 81960
ggactagggc tcgattgtta tttatttatt tatttattta ttgtgtgtgt gtgtatgtga 82020
gagagagaga gagaaagaga gagagacgag gtcttgctcc gttgcccagg cgggagtgca 82080
gtggcatgat catagcttac tgcagcctca aactcctggg ctccagggag cctcctgcct 82140
cagcctgcca agaagctggg ggtactgttg tgtgctacca tgcccagcta gtttttaaag 82200
tatttttttt tgtacagatg gggtcttgct ttgttgccca ggctgggctt gaactactgg 82260
cttcaagtaa tcctcctacc tcagcctccc aaagtgctgg gattatacat atgagccact 82320
gttcctggtc tagtttgcac ttttttttttt tttgaaacag agtcttactc tcttgcccag 82380

gctggaatgc agtggcacag tctcagctca ctggaaactc cacctcccag attcaagcga 82440
ttctcattac tcagcctccc gaatagctgg gattacaggc acccgccacc acacccagct 82500
aatttttgta tttttagtag agatggggtt tcgccctgtt ggccaggctg gtctcgaact 82560
cctgacctca ggtgatccac ccacctcggg ctcccaaggt gctgggatta caggtgtgag 82620
ccaacacgcc tggccgtttg cattttaag ataaaaattt taccatgctg gatatattgt 82680
agtagctatg tacttcagtt tctcaattgt taaacgaact taatagaagt actaccttat 82740
agaattttgg ggattaaatg aaatagtctt ctgagcacaa caaatatatt atcagcacaa 82800
cagctagctc tattgagtct ttattattat aatagcagta atagtcagac ttggaagggg 82860
tgaaagagaa caacagtcca ttttatttt gtggcatata tatcataggt cgtaagacct 82920
tggattgttt agatgccatg ttaaaacttg acaaaactag aaatgttgtg agtgtgcaat 82980
agcaggtgat aactaatcca atcattaatt tattcttgaa tttgatcaga agacagacct 83040
aacttcatct attgccaatt actattgtaa caaaatctat tggaatttca gtttaggcac 83100
tgcagtacaa cagtgtgaat ttcaaaagtg agatatttta tgtggctttt taaagtaggt 83160
tttcaaacca gttaaaggct ctaaaaccca ttaagaagga tactatgggt caggaatagc 83220
attttatgag gactcttgag aaaacggtgc tgtgtcccct ccaccatggt acaagtggag 83280
gttactctca cttgttctct tacttggtat ctctccttgt tctctcactt ggtatcgagt 83340
gagaggggct tgagagggac gacttaagag atttaaagga ttccccccag ttgtggggaa 83400
cagaggacca gggtctgact ctccgagaca cccaacaagg agagaactgt ggctggctta 83460
ctgttcccac agaaagagct atactgtaat ggtgtctgtg ttgaatttgc ctacactcca 83520
gcctttgtca gggcaaagga aatgtatttc ctgcagatta gaggtgggcc tagcggagga 83580
gaagggtagc ctggagtttg actcctgtcc aaatagatgc ctgaaagggc agtgaggttc 83640
caacggtgag tcttttgaat gacagccaaa aagccaggag aagaatgaat tatccaagtc 83700
aatggtgcta tggccagaag gaactctgag gtgagactct gaataaccca tgaaagtgca 83760
tctgagaaaa aagaattagc ttcaaacatc agcagacagg tgttttggca ggagggctca 83820
tgttggaatc tccttgcctt ccttccccag cctccatctc ccatgcctac cttggaaaga 83880
tccgacagct agggttggtg taaagtcgtg gaagagaaga aaagcagctg atttcaatcc 83940
cttcccaggt tccctggggc tggggaaaag atttgattac ctagtgaata ttggtttgtt 84000
acctaaagtg accataagcc tttctttaac attgaccaaa agaatcaaat gggcctgttg 84060
aagtgttcat ctagtgtcaa gggaaaattt ttccccactg aataaatttt aagaaggcag 84120
tcaagacaag aagctatatt tgattatatc ctgttagtgc ttattcaata gacacataaa 84180
tctgtaattt ttaatatttg gtatagaagt aggttgaaat ccacagtaat tcacagaaac 84240
ttgtgcaagg gttttgtttt cttttctttt ctttcttttt ttttttttt tttgagacag 84300
aatctcactc tgtcccccag gttggagtac agtaggatga cctcggctca ctgcagcctc 84360
cacctgccaa ggttcaagca attcttgtgc ctcagcctcc tgagtagctg ggattacagg 84420
catgagccaa cacgggcggc taatttttgt atttttagta gagacgaggt gtctccatgt 84480
tggccaggct ggtcttgatc ctgacctcag gtgatctgcc tgccttgatc tcccaaagtg 84540
ctgggattac aggtgtgagc caccatgccc ggccaaggtc ttttttcttg aaaatatctt 84600
cactcatata agcagtatat gcaatataag gatatgctct tgggtttctt gatgtggtct 84660
aatatttagt gttggcccct taattataaa agttgctttt atctaaaagt agatgttagt 84720
tgtcaggcaa tgtttgctgt aaaaaataaa taaaaaataa aagtagatgt taggatattt 84780
tcttctagcc tgcctagtat ttatattaga ttctttcttt ttttgagaaa gcgtctcgct 84840
cagctgccca ggctggagtg cgcagtggcg cgatctcggc tcactgcaac caccatctcc 84900
cgggttcaag tgattctccc atctcagcct cctgagtagc tgggattaca ggcacccacc 84960
accacgcctg gctaattttt gtattttagt agaggcggag tttcaccatg ttggccaggc 85020
tggtctggaa ctcctgacct caggtgatca gcccacctcg gcctcccaaa gtgctaggat 85080
tacagatgtg agccaccgca cccagcctag attgtttctt aaaccataga tgtctgaact 85140
tttttgaatg aaattaaatg attagagatt agtaaaattt ttgtataaga tagtagacta 85200
acaaattctt actagtctgg gtgtggtgac tcatgcctgt aatgccagca atttgggagg 85260
ccagggtggg ccgaccactt aagcccagga gtttgagacc agcctgggca acatggcgaa 85320
accttgtctc tacaaaagat acaaaaatta gctggtgtgg tggcacacac ctgtagtccc 85380
agctactcag gaggcttagg tgggaggatg gcttgagcct agaaggcaga ggttgcagtg 85440
agcaaacatt gcgtcaccgc actccagcct gggtgacaca gcgagaccct gtctcattaa 85500
aaaaagaaaa tttaacctgt ctcaagctct ccatactgta aggctctgca tgtcttgatt 85560
ggattgtgct aatatatttg gccaatcagc tccttcctac tgtctacttt tgaatccctg 85620
tcaccaccat ctaagtcaag atgacagtgt ttacacagtc tcttcattgt gttttaagat 85680
tatagtcttc tttctggtga gggaagaaag aaaatagaaa tatggcttac tgattgggcc 85740
atggcttacg cctgcaatcc cagcatttta gaggccaagg tgggaggatt gctggaagcc 85800
aggagttcaa gaccagcttg agtagcaaag tgagaccctg tctgtacaaa agaaacacac 85860
acaaaagaaa tatgactgac taaaatacat ataattttca taatacttta aaatgtaaga 85920
aggcaaaaaa tttctgggct caaggtgggt gatcgcttga acctaggagt tcaagaccag 85980

```
cctgggcaac ctggcaaaac cttgtttcta caaaaagtac aaaaattagc caggcatggt   86040
ggtgcacacc tgtggttcta gctacttgga agattgaggt gggaaatttg cttgagcctg   86100
ggctgtcgag atcacagtga gctgagattg caccactgca ctccagcctg ggcagcggag   86160
tgagaccttt tctcaaaaaa aaaaaaaaaa aaaggcaaaa aattaaatta ttagtatggt   86220
aaagtttcgt ttggacttaa tatgaaactc atttctagaa atgatgatca tttgcatagg   86280
gcttaacttc ctttgctaag aaaatagagt agtatactag gagacttcca gagctgcata   86340
gagcttcagg gtcatctacc aagacagaca atttgttgtc atcatcagtg ttaaactcta   86400
aattattaag tgcttatgtg ccagatactg aagtttatat acactttctc taatctttaa   86460
taattctaga aaggtatgtg tttgatccat tttcaagata agaaaactca ttcagagggg   86520
aagtaacctg accaagaata cgttgtagtt gtagagctgg gaatatgact catgtctgtt   86580
ttttccatag cccattttca cttaggttcc attagtttat tatatatttt agaaggactt   86640
acagacttga tttcctcagc gaagtcagat tactcatttc tttggcacag aattatgatt   86700
atttgtattt aatatttctt tcctgttggt tttcattttt gagtttttat ggaatctgac   86760
ttctgattct caatttttt attgtgaaat ataacatacc ttcagaaaaa aagattaaac   86820
atatctatgg ttttataaat gattataaaa taaataccca gtaactatta tccaggtcag   86880
caaatattct actagtgtat gagtcaattt ccatggcaaa agaactaagc ttaggcacta   86940
tactcaaaaa aactaaaaat aaaatttttc taaatgtgta ttatatcaat ggaataaata   87000
caaatataac ttaccatgtc ataattcccc ccacgctttc ccttctttta cagcatgggt   87060
aggttctctc tccatgggga tgattttctt ttgctgccca atagtcagcg tcttcacaga   87120
cctatttggt tgtcggaaaa cagctgtcgt gggtgctgct gttggatttg ttgggctcat   87180
gtccagttct tttgtaaggt aaggacttgg ttttttcatg ttgcttttta aaaactgtta   87240
gataccttaa agttttactt tcagaaacta tgctatttac aagcaaagat cctccttttc   87300
atttttaaa actttaagca atatgactta taaaacaaac tgttatccat agcagcaaat   87360
cagagcttga gaatttgaat gctttttttt cttgtaatgc ctaagactta gacatcaatc   87420
aggatatatg tgtttcttgg tgcatggaaa aatgtttctc cttatctttt tcttctattc   87480
acataaaaat cctaatggca ctcaagttat aacaatgatt tatctaaaac agcagcctta   87540
gtttagggtc agttcagtct gaggcccacg ggttacaata agtggtgttt taaagtagct   87600
gttcataggc ttgatatgaa atatttttgt taatgagacc aaaactttgc catttattcc   87660
aacccaggta gagaattcct gtctgttctt taaaaaaaga acatgctaaa attttaaaat   87720
atcatggcaa aatgaagtgg tccaatgtac cttaaaataa aacttaatgt caatgtactt   87780
ctcctgtatc tattagaata aggatcccca acccctgtgc cacagactgg tacgggtcca   87840
tggcctgtta ggaattgggc tgcacagcag gaggtgagct gtgggtgagt aagcaaagct   87900
tcctctatat ttatagctgc tccccattac ttgcattacc acctgagctc cacctcctgt   87960
cagatcagtg gcagcattag attcttctag gagtgcaaac cctattgtga actgtacata   88020
tgagggatct aggttgcact ccccttatga gaatctaaat gcctgatgat ctgtcacttt   88080
ctcccgtcac acccagatgt gactgtctag tttcaggaaa acaagcccag gattcccact   88140
gattctacat tatgatgagt tgtataatta tttcattata tattacaatg taataataat   88200
agaaataaag tgcacaataa atgtaatgca cttgaatcat ccctaaccgc cctcccccgg   88260
caccatccat ggaaaaattc ttcctgaaac tggtctctgg tgccaaaaag attggggatt   88320
cctgtattag aagaaaagac acagtgcact gaagagggac tcactgtaag gaacggatgg   88380
gcacataact gcatggaaca tctctccctg cagctctagt tctccttgta agtcccttgc   88440
gttggtagaa taatcctcag ttagacaaac actgatttaa tatgtagctc tggctaacag   88500
gaggtgatta agaagaaaac ctcttaagat gatttccatc ctttgtctct actttagtgg   88560
tttatcttca tttcctgcct ctttcctgtc cctagctgtc tttatgctgc tttagttgaa   88620
aagcgttaat gtggtcatta aggaaaaata agtcaaattt acatttgact ttttatttt   88680
aaatatttaa tcaacagaat ccttggtttt actcattgcc gccccccacc ccccaacaca   88740
catcccttcc tcaactctaa agtgagcctc attctttcat attttcttcc atctaattta   88800
gaaattctat ttggattttt aaaaattata tttatttctt tgtagaaaat agatattttc   88860
atctttaaag taccttttatg ggtttttttct tctaaaattg ttttttaaag aaaaaagttt   88920
tatttggaat aagattcttg taggtaattc catgagatga tttattttag cagcaacata   88980
atatttacat tattattaat gtaattaatg ttattaatac ctcatcagat agcttctttg   89040
atctggaagc ttccaggtac ctattgtcag tacttgtggc tctaccactt gccgaatgta   89100
ttacaactct agttgtggta gagaggggac tgagaggtag acaacttatg taatctacta   89160
cctagtttgt taacaaaaca cacatacaaa gcaatgtttt tcaaattttt ctgaccactg   89220
agcaataaaa attatgacat atattttgat gtgacccagt tctgtctctc tttctctacc   89280
ctctaagtga aacaaaattt attgaaacca aaattccctt actacatgta atattctcat   89340
atattctatt aaatttcgtt atttagcttg ctgatcaaag gctactgaaa cttgagagca   89400
agatacagga gcaaggggaa atgtggtata gattctgagt gtcaagtggc aggtccattt   89460
tttcctctag ctccagttct gccttctgag gaaaaccttc tccaacaact taggtcaatc   89520
acacccatgt cccttctctg aatccttttt gcacatatga ttggtatccg acagccttac   89580
```

tcatttacat tgcacttatt tggctgccaa acgtcacaaa ctggaaccat gtgttactga 89640
agggaaaacc tggaagtgaa aagggttcag cagtagtgca aataccatca taaagctcat 89700
atacttcact ctgcaggagg gagaagctct gtggttttcc aactgagagc attacagtac 89760
agtgatacca ctgtacagga actgatgttc ctgatgattc tgctgtgaac agtattttta 89820
atatacactt tgaagaaggc agagagaaat gtataataga cttaaatttt tttctttaaa 89880
attgttaaat aaaaacaaat aagcacttta agtaagttac aattatctgg aaaactactt 89940
aggtggaaaa actgatacag aatgaatgaa gtattaattt ctgttttgtt ctgtgttatt 90000
attatttggg atagatgtct tgtttcttta agcagactat gaatatcttg aaggcagaac 90060
cacatttttt ttttttttga gacagggtct cactattact caggccagaa tgcagtggtg 90120
ttatcatagc tgactgcagc ctggattcct gggttcaagc cgtcctcctg ccccagcttc 90180
ctgagtagct aggactacag gcatgtgcca tcacacccag ctaatttcag ctatttttt 90240
ttttttaaa tagagatggg gttttgctat gttgccagac tggtctcaag ccatcctcct 90300
gccttggcca cccaaagtgt tgggattaca ggtgtgagcc accacgtctg gccaaggacc 90360
agatttttaa tattcttttc cacaatgtat ctggtacaca gtagttgctt aatatgttgg 90420
ctaaacaaag agtggagatt cagtaaaggg tgatcagagt gaggtgagat taatttggga 90480
aagcctagaa gtgattcttg agcctgattt gaaggtggtg ctagctgtgg attaatagag 90540
ggagaagggc atctcagaga gaggattgcc aacatgcctt aattttatca gattctagag 90600
ttccttatga ttacctcagc atgttgctag actagcatta ttatccaaaa ttttaattat 90660
taaccaactt taatcttact ttctaacaaa ttgtttgctt ttactactga tagccttttc 90720
aaaaaacttt aactagtttt attccttacc ataattgttt caaagaacat aatgatatga 90780
tcctttatct tcctaagaaa tgtgcaatta tttggttaaa ctgtaagatt atttaatcca 90840
ttattctttt gacacatgca tggccttaca gcttacaaac tgggatcact aaaggaatac 90900
acttaattta agtctttctg tagtcagaat atgatttctt gttgtcttgc acaatactga 90960
gaacagtgca gtacagggcg aaggttggtc tacagccctt aggccagcaa aaacaggcac 91020
aactgcacct ctgtgcaaat gttcctgaca taaccttggg gaaaaaatat aaaatgcggc 91080
cttttctttt actaccttgt ttggtaagta cctggaaaaa ctccatgaaa taattagatt 91140
tcatagttaa ttctaacttt tttaaaaaat gtttcattga gactaggttt ttggtttgtt 91200
aattgaatca ctgttgattt taccccttcct ggcaccaacc tttatttctg agctgtggag 91260
agcacagttc tcactcagtg ctgtgtgcgt cacctgaaat ccacagaaag aggtggctga 91320
acaaaatcac tgatgacctt aatggttatt tttcacatat tcagattaaa ttaaaatacg 91380
tttagtgcta catgcttgac ttactgagtt tttccctcta ttttggttaa ttttttttt 91440
ttttggttaa cttttacttg tagaaaatat gttgatgaac aaaaacccac ttatactata 91500
agattttatt ctaccaagca cacagtaaca atattgaaag ctgctttcca tctttttcat 91560
ctttatacag ttccatcgag cctctgtacc ttacctatgg aatcatattt gcctgcggct 91620
gctcctttgc ataccagcct tcattggtca ttttgggaca ctatttcaag aagcgccttg 91680
gactggtgaa tggcattgtc actgctggca gcagtgtctt cacaatcctg ctgcctttgc 91740
tcttaagggt tctgattgac agcgtgggcc tcttttacac attgagggtg ctctgcatct 91800
tcatgtttgt tctctttctg gctggcttta cttaccgacc tcttgctacc agtaccaaag 91860
ataaagagag tggaggtagc ggatcctccc tcttttccag gaaaaagttc agtcctccaa 91920
aaaaaatttt caattttgcc atcttcaagg tgacagctta tgcagtgtgg gcagttggaa 91980
taccacttgc actttttgga tactttgtgc cttatgttca cttggtgagt atgctccttc 92040
actgatcatg aatattacta tttaataaag aaaaacttct ttgaagagaa agttaggtcg 92100
agttaaagtt ggcctcaaac attatcctgg ttgtaatttt ggtattcttg aaatgaaagg 92160
tctctcaaga caatgtcagc acatccatta gaccactaaa cagagagagt atgtttcata 92220
gtgtgctttg gtattttaaa aaccctgcaa acccagccag acaccatggt gcctgtctat 92280
ggtcccagct actaagctga ggcaggagga tcacttgagc ccaggagttc gaatccagcc 92340
tagacaacat agagagactc tacctctaaa aataaaataa atgtccccaa acaaacacaa 92400
tgtttttaa caggaaggct aaaatagtgg aacaaattac aatcagtata aaacatttga 92460
taggtctctt tttcttcata tggcttttat cagggacaaa gctagcgcta tgattttgct 92520
accataagta aattgttttt caaccgaagg gtgtaggtaa ttagcaaaaa agccatgatg 92580
ttgatacaaa gaaacattac atctacttgt ggtacacttc tgggaaaatg ggaattctat 92640
tcagaggaat atctgagaaa agttactcaa gatctaaatg aggaaagaga actatggttt 92700
tataggaaat taggatttca agtgctcaag aagtttatat tgtttatttc tatttcaaag 92760
gcaaaattca gctttgttat actgaaatac gaataattaa tgtctagact ggggttggtg 92820
cctcacgcct gtaatcccac actttgggg ggctgatgca ggagttcaag accagcgtgg 92880
gcaacataag gagacttcat ccctacctgg ggaaggaaaa aaaaaaaaga aggaagaagc 92940
agtgtctaaa gtatctgccc ctggcaacgt ttgttcaaaa gtgttcatta tgtttcttcc 93000
ttttttcttt tgtggctgaa aatgtattta caattcaccg taaatgataa aaatggcatt 93060
ggcacacata tttgtatgtt tgtgaacttg gatttttttc tagcttacag tctacttttg 93120
gagatttgtg caatttttct ttagttaaga aataagtata aatataaccg atttacggac 93180

tatcaggcta catcctgatc tgatagtcca ttttcatact attaggaaag tatagccgaa 93240
ccaacttaag gtaagtttcc tggaatatag atctgttgtg acaggattaa ctttaccatc 93300
caacctcttt catagcttct gtagtcaaga gaacatttat tgtgtccttt cttaaaaaga 93360
tgagtagaaa ttctttttct ttttttcttt ttttccagac agggtcttgt taagttgctc 93420
aggctggctt caagcaaccc tcctgcctca gctaggatta caggtgcaag ccaccacacc 93480
cagctttaaa aaaaaaattc tctttggtac taccacatga acacacctag agaaatcata 93540
actcagcttt gctaatacta gacatttacc aaaggaaaag tggtagatga ctgtctagtt 93600
atttttggtt atatatttat aatttgtaaa ttaatttcac atatattact tcatttgact 93660
ttcacaataa accagtaaag cagataaaat aaatattagc tccaatttta cagactgaaa 93720
aacagatcta ttgttaatag agacgttaag tgattttcca agaattacat gtcagtaaac 93780
agcagagcag gagttagttc tcttccactg tgcttacctg gtagcaaaat cagtctacag 93840
tcttaatagc atattgggcc acttccctgg atatattacc aaatgtgtcc atcttattag 93900
gggaaaaatg agtatgccta aggaaattta ataagcatgt tatttcttca ggtaaataaa 93960
atttcatagt ggaaggtgag ttagacaatg ttatagatac ttttgtgatc aggagatggc 94020
aaatcagatg gtgcacagaa caataaagtc tctgttaatt ctgttaataa accatgcctt 94080
ttttctgctt tcccttcttc caggcatgtt ttcttacaaa atatgttgac attgttcatt 94140
tgagattttc tctttctcat aacggtgccc gttatcgcac cgaatgcagc acggtagagg 94200
aaagatcaga tagctaaatg ccatacaggt gtttaaatct cctctttggt tatgtactga 94260
gtttgtcact ttgttgtaat ttaaggtttg aattatggat acttaaccag gaatgggaca 94320
ctagtttcct ccttatacag ggaaaaggtg tctcatatcc ttcaaaagac tagtaaagta 94380
gatgatgttc aattcctact aaacccttta ttgactgttg aggggacaca tatatgagac 94440
gtaaaaattt gctctgaagg agcataaacc tagtacatgt aattaaaaat ggctacagtt 94500
tataaagcac ttttacatac attctcttat ttaatattca caacaatgca gtacctgtgg 94560
tgtatcctct ttatttcatg gaagggaaga ctaaggcccg gaaagattaa ataacttgct 94620
cagccaggca cagtggctca cgcctataat ctcaccactt tgggagaatg aagtggaagg 94680
atcacttgag cccagcagtt caagaccagc ttgagcaaca tagtgagatt ccatctctac 94740
aaaaagtaat ttaaaaaaat tatctgggca tggtggtgca tgcctgtggt cccagctact 94800
tgggaggctg gggtggaagg atcgcatgag cccaggaggt caaggctgca gtgagccatg 94860
atggtgcgac tgcactccag cctgggtgac tgagtaagac cctatctcta aaaaaaatta 94920
taaagtattc taaaggaaga acagattgaa caatttttaa tttatttgtc tcctcctcct 94980
agtggcagcc ttttaaatat ggaaggtgaa gaaataaaga gccagatgtg gtggtacaca 95040
tctgtagtcc taactactca ggaggctgag gcaggaggat tgctggagcc caggagttca 95100
aggctgtggt gtgctatgat tgtgccactg cacgccagcc tgggtaacag agcaagactc 95160
tgtctctaaa aaacagataa taaataaaga agtaacttgc ttgaggtcac agagatagtg 95220
actgataatt attactgtag tacttttatg taagaggcag tattgtatag tggtttaaaa 95280
gtgaaggttc tgggcctggt gcggtggctc acccctgtaa tcccagcact ttgggaggcc 95340
aaggcaggtg tatcaccaga ggtcaggaat ttgtgaccag cctggccaac atggtgaaac 95400
cctgtctcta ctaaaagtac aaaaattagc tggacgtgat tgcttgcacc tgtaatctca 95460
gctactcagg aggctgaggc aggagaatcg cttgaacttg ggaggcagag gttgcagtga 95520
gctgagaccg cgccattgca ctccagcctg ggtgacaaga gcgaaactcc atctcaaaaa 95580
aaaaaaaaaa aaaagttctg aagtaagaca gatctggatt taaattcagg ttttgtttct 95640
tactagttgc ataaccttgg gcatcctctg taagcatcag tttcctcatc tatggagata 95700
aacccaattt tgcagagttg tgaggattag ataaaatgta tgtgaaacat ctacctcagt 95760
tctggcataa aaatgggagt tattttaatg taaggcaatg tgattgccaa cttgagatag 95820
aagtaaattt tgaaaggaga aagataatac ccatttggaa aagtggtttt aaaaagtttc 95880
atagcattgg agttgggcct tgagcatgag attttgtgta caaatctgat ctttgatcaa 95940
ctagggaact aacttaccag tttaggtctt tgaagattca gaaatacaat ggagtgctct 96000
cattgctatg ttaaaaattc taagatctta ttagattgta catgatgatt tgagagagaa 96060
tatgtatgct tgctttcaaa gtgaggttgg aggtttgatc ttctcgtagt tgacgtttca 96120
aaaagaagaa ttagattgcc tcctcgaagc taaatttacc tttcttttag gccttcccac 96180
ttaaaatctt ttttagaagg atacaaatct tatagatcaa tttagatgag gcctaacttt 96240
ctaaaaacga ttcctagtag cagctgcatc agtttttatg aattgcccct tttgcctgag 96300
agttgttttg ttttgttttc tggaatcttt ttttgttttg ttttgttttg ctttgttttt 96360
gttttcgttt ttttttgaga cggagtcttg ctctgtctcc caggctggag tgcagtggtg 96420
caatcccggc tcactgcaac ctctacttcc cggattcaag tgattctcct gcctcaacct 96480
ccctagtagc tgggattaca ggcgcctgcc accacacctg acttaatttt ttgtatttt 96540
agtagagaca gggttttgcc acattggcca ggctggtccc gaactcctga cctcaggtga 96600
tccacccatc ttggcctccc aaaatgctgg gattacgggt gtgagccacc acgcctggcc 96660
tctgggtttc ttttttttt ttttttttt ttttcttttt aacggctcct ctgactcctc 96720
tcatttagct ttcaggagca taaactctct tggttttctg cctacctcca catcactcct 96780

```
ccttagtttc tttgctcact tcttcttttt cccactgacc cctgaatatc agcatgtcct  96840
agggcttgtc ccctgatctt tttctccatg tattctactg gtggtttcat ccagtctcct  96900
aagttcatac atcacgtata tgtcaatgac ttcaaattta taattctggt ccagaccttt  96960
tccctgaatc ctccaccaga gctgtatatc cagctgctta cttaacatct ccacttgggt  97020
aactgctagg tgtttcagac ttaccctgtc taaccctgag gtcttgatct taccccttaa  97080
aacttactct gcccccagcc atcctcatct caggagctgg caattccgcc ctttcagttg  97140
atcagactca aaactttgga gtcatccttg gctcttcttt cttgcacacc atagtcctga  97200
tccagtgaag aaatcctggt ggcttttctt tcaaaatata tccaggatct gaccacctct  97260
caccatcctc actactcata ccttagccca ggctaccacg taccctagc ctggatcact  97320
gccagagcct cctaactggt ctctctgttc cttctctgcc ccgcggagtt tgttctctat  97380
gaagaagcca caggcattct ttctaaacat aagtcactct gctcagaatc cttcaatggc  97440
ttcccatttc cctaagagta aaaaccaata tccttacagt gacctacaag gtccttcaca  97500
atctggcccc cactacctct ccgagcttcc atcgctgtcc cttgcccact ctgcttctgc  97560
cattcgcctt ttaatggggc tcactctgac tacctgcttg aaacttcctg cgtccctttt  97620
cccctgagta ttcacaaacc gctcctagta ctccttttct ttttttgtag cacttaatac  97680
tttctaacat tatctatttt acttctttat tgtagtcatt gcttactatc cgtatattta  97740
cacgtctgct agaatgtaaa caccacaagg gtaaggatct atttcattca gtggtagatc  97800
ccaagcatct agcacagtgc ctagcacaca ctgggtgctc aaatatttgt tgaatgacta  97860
aatatattct gggtgagtct gaagtgacac tgtataagta atgttcattt tttcatcatt  97920
tggatcttta aaattctcta ctttgatgct ataatgattt ttcacattct gtacttgcag  97980
gacatggtgt tattaatatt tattcaatac ttattcaaca aataagctca aactaaggaa  98040
acctcggaat aattgagtaa ccagtaatgc tgtccgttga tggaggagag agttggtgtg  98100
ttttgctctg attcacttat gcctttgctg aaatttttaag ataaatagaa gaaatttctg  98160
gtccctcaag taactgtgtc ttcagtaccc actgaaaaat ctcaaagagt ctggagtggt  98220
gtgtttaaga ataggatgca ggatgcagaa ccataaccag gcctcaggtc tgcatagctt  98280
tggtcgagca ttgagcatag ggcctcgtga gataactgat aaatgccaaa tatgacaatg  98340
ataaatgcca aatatgacaa tgataaatgc cgaagaatga cagtgacaat gataatgaag  98400
ttaccaaaaa tgatggtaac ttttctcatt ggcatgaaat gctctatctc caatctgaag  98460
ctgatgatgt agtttcagtt actctcatct ctctcccctg ctactcagat tgaaaatcag  98520
ctacttagta cctgtgttct ttgactctag accatatcat tgggtcaaat ttcagttttt  98580
aaattttaga tccacatggt tctctgtcaa gaagatgact gactcatatt gaaatctgta  98640
aaatatgtat tcattagcct gtttttaaa aactccctta taagtgggtt gactttgtgg  98700
cagatagtaa ttgactgttc tcaaaagaaa ctttgacctg gtaggaagat cccatttacc  98760
tgatgctatg gttcaagaca gacagatcat ttgcttgcta gcagggcaat taggtgaact  98820
tcaagtccac tagtaattgg aaatgatttt tttttttttt tgagactgag tctcattctg  98880
tcgcccaggc tggagtgcag cggcatgctc tcggctcact gcaaccttca cctcctgggt  98940
tcaagcgatt cttctgcctc agcctcccga gtagctggga ttacaggtac ctgccaccac  99000
gcctgactaa tttttgtatt tttggtagag atgggtttca ccatgttggc caggttggtc  99060
tcaaactcct gacctcaggt gatctgtctg cctcggcctc ccaaagtgct gggttatagg  99120
cattaaccac cgcccctggc catgaattgt atttttaaac cagaaatgaa aatttgagac  99180
taataagtca gtacagggag catgtaaacc tcgaaaggta ttttttagct ttgagtagtg  99240
ccagatgctg ccaagggtca atcaacactg gaatgtagct attagacctt gctaggcaga  99300
gcacctccat ttacactgtg gtcagagcag cagtactgct ccaagccaga gctaagggcg  99360
ctgagccacg caaataggaa cagcatacaa gccttcatct ctctgtggct tcctcagagg  99420
gagattcatg taacatttgc caagaattga ttatgtgtca agcacttccc caaaatctca  99480
cagaaccacc acaaggatga gtgtaataaa taacacatac ttagagccaa ggaaacaatt  99540
ctgcaaagct gtgcttgttc aaagccattc gcattatgct taaagctggg atttgaacac  99600
aggtttcaga caaatatgtc tgaaatatac tctttttatg aaggagtctg cattccttca  99660
ttgctaatcc agagatagga gtgctgctat tttcagccat actgggccta caccaaagat  99720
tgctttgcac gtttcccttc tgttctctca gaacgaagaa cagaggccat gttgagctgt  99780
tccagcgctc agagcatgct tcacagccag ggagaaaact ctggaggaaa ccagcttttg  99840
ttttgatata attaatggga atgagaaaat atctataccc ttattttcag ccccaacttc  99900
tcttttgatc tcaagtacat tgtgaatatg agaaaactga ggccatgcag ttacttttca  99960
caacctgtga caagcagaac atggaccata catagctttg tgttcaattt tgctttctac 100020
agtaaacatt aagcataaca gaagaacaaa aatggacatg tacaaattta tagcaagatc 100080
tatcctttat ttgattaaca taaatactat tgcaggaaaa tggaaaaagg taaactgctt 100140
gaaatttagt cacatataaa cgctccgagg ccactggtgg atcattagtc tcctgagaga 100200
gctctaaaga attagtgtgt tggaaaactg ttccctcctg ttaatgtgta aatttcacca 100260
gtgggttttt ttttttttta agacagggtc tcactctctt gtccaggctg gtatgcaggg 100320
gtgcaatcac acctcactgt agcttcgacc tcccggactc aagcaatcct cccacctcag 100380
```

cctcccaagt agctaggacc acaggtgcac accaccacac atggctaatt tttaattttt 100440
tgtagagatg gggttctcac cacattgccc aggctggtct caaacttctg tgctcaaaca 100500
atcctcctgc cttggctccc tgccaaagca ctgggattac aggtgcaagc cactacatct 100560
ggccttcact aggtttttca ttttgttttg catgtgtctc aggttttatt tgataaaatg 100620
cagtatactt tgaatcatct caaatttcat ttctaatatg gacattggca tgtctcaaat 100680
ccttggacta ataatcaaat taaagtttgt tcaagtttga ggaacctaaa ttagccaatt 100740
agataagggt cctttcatgt ttttatatca actagaaaat aaattgtttt gatatgggat 100800
gaatagaaat agaaatctta atttgaagaa tcttcccctt gtgaggctat actaaatggc 100860
ttttgcctga tatttacaag gtggctttgg gttgtggaga gagttgtctg atccattgag 100920
agtacatttc ttaccttcaa catctagggc atcctttggg agaagccctt gtagtcacta 100980
actctaagga tcatagagca taagggtaag caggccttct tatgtattca tgctatcagg 101040
aagggtcttt agcacccaaa caaagttcta ggggctgtac attgctgatg tgttaaccct 101100
cagctgccca tgtagcatct atttacccct atgctttccc cactttctat ccctatcatt 101160
atatctctgg ctcttttgcc cctctctcct tgggcagctt acttgtaatt agaaagttta 101220
tattccctca taacatattg taaaagtgct catttaaagg gcaatgcaca ccaaattgga 101280
ggtgtataat tgcaaacatg gaatccctat atctctgtta tgcaatccct gtatctctgt 101340
atccatgtta aattgaactg atgcttttt gaagtaaaat ggtaagaaca gtggcaacat 101400
ctagtcttca gagcatagtt taagattttt gcccaatcct ccaacccatg caatggtgtg 101460
ctttgaaaac cacaggtttc ttttagacaa atacaacatt tatttcccgc atttcttttt 101520
gatttaacat tttagttaac atttttatta acattttagt ctacaagatg cttccacatt 101580
atctctcatt ggagtctcag gaccactgtg tgaaatgggc aatatcaggg cttctatcta 101640
gcaagaaaag aaccagattt ggggtggcga aacaacttgc tcagggttgc aaggatggta 101700
catggtgcag ccagggcttg agcttgggtc ttcttaaaag tgtggctttt aaataaaata 101760
cttaagtgcc tgccaaaaaa gtataacatt aacttaggac ctgaaaggca ttgtacagat 101820
caggtagttg cactcctccc cctgccctac aaaàaaagaa aggtaaagga acgaaggcat 101880
ggaatagtta agttgcttgc caaaagccac agttataaaa gtagcagaac tgggtgtaat 101940
acccaagaac atccatggaa aataaatgga agcttattac agcccagcct gtaaatatgt 102000
acatagaaac agaatgtgta tgtagaaaca aaattattag aggagtgaaa ttagtttctg 102060
tcctaaccct ggtcaactca tagggttcat ttcccaatca ggaggttgct ttgccatttt 102120
gacatgatcc attttctcag agttgccaag agcaattgca actggctttg ttgtctttct 102180
ttcatggatc tatttggagc tcagggacaa ggttgttatg cggcagattg tccctgacaa 102240
tgcaaggctc acagtagtat ccttataaac caaagttcac tgtacactgg gccaggaatg 102300
ggctgctgaa ggtgacctct ccttcatgtt ggttggttta tctttgttgg ttggttgtgt 102360
ttatctctgt tcactataag gttctgacag aagcaaagtc ttggtccccg ttacatgtcc 102420
ccagcctggt ggatgatggt tacagagtgc tggaaacttt tttatttatt taaaatggag 102480
ttacgtaact gaagagttgc cctatctcag tcagcactgc aaccaattaa aatagaaagg 102540
cttaaaatat taattttgtt tttagcctag agtcttaaat actagggttt aaaagtttca 102600
tttactttct ctctccctct ctccctctct ccctctctcc ctcccttcct tctttcttcc 102660
ctccctccct tctttccttt ttccttttte ttaactatgc agaatgcttt gctactttgc 102720
agaatgcttt gctactttgc aaaataaaga gtaatcactt tgctttttct tcaagacttt 102780
caaatttaag taatttgtt tttctatttt tcattcaaaa atagggttat aacatttcct 102840
tgacacaagg atatatatat atatttagag acagggtctt gctatgttgc ttaggctggc 102900
cttgaactcc tggtgtcaag tgatcctcct gtctcagcct tcccaagtgc tgggatcaca 102960
ggcatgagcc actgtactca gcctttatta atctaaatat gataatttac ccactgagat 103020
tcatttgtgc tgattaaatt tactctcaat ccctatatgt atttcttgta tttttgttgt 103080
tgttatgtgg cctggaaatg tttcctttat tcattcattc attcccttcg caattaactt 103140
ccaaaaaggc tatgaagata tttatgcaca tatacatttt atgattcaac tctcatggta 103200
acctttttata aggaaagcag tgctaatagt ggttcttcta ctaatagata agaacattga 103260
ggctcagaga ggtttaaaag gtttgcctaa ggttgcacag ttaagtaaca gagttgccat 103320
tagaatagat ttccagttta ttctaatggc aaccttaatt attgttgttt gacttgctct 103380
aattataaat cagttaaaca gattcaggca ttcatttgag tatctattgt gtgccataga 103440
cattctagat gtggggatac aatgctaaga acagatactt tctcatacac agctcaataa 103500
acaagtaaat tcataaagaa acaaagtagt tttcaaatag tgatattttc ttcaaaggaa 103560
acaaattgag atggaaggtg atggtaagag atggtaagag acaactttgg ccagtcaggg 103620
acaacctcat tgaaaagctg atagtaagta catcctttgg gtaaagggta gtataaggta 103680
ctttgaaggt acaaaaataa gacagctttc tattgccctt gggaggccta taacagaatt 103740
tctcaagtct ctaaggccaa tcaagagttg gatttttta tccaacttat ttttaattga 103800
tgtattatta aaaatctgca tatcaaaaat gaaaatgtct tgcatacttt gctgtaggac 103860
ccaatcattg ttttttcttc atatactgca ttaatctgtt ttcacactgc taataaagac 103920
ttacctgaga ccaggtaatt taggaagaaa aagaggttta atggacttaa agttccacat 103980

```
ggctgggtag gcttcacagt catggtggaa gatggaggag gatcaaaggc atgtcttaca  104040
tggtggcagg caggggagta tgtgcagggg aactgccctt tataaaacca tcagatcaca  104100
tgagacttat tcactgtcac gagaatagca caagaaaaac ctgtccccat gatttaatta  104160
cctcccacca gcttgctccc atgatatgtg gggattatgg gagctacaat tcaagatgaa  104220
atttgggtag ggaacacagc caaaccatat cattctgccc ctggctcctc ccaaatttca  104280
tggcctcaca tttcaaaact aattatgcct tcctaatagt ccctcaaagt cttaactcat  104340
ttcagcatta actcaaaagt ccacagtcca aagtctcatc tgagacaagg caagtcccct  104400
ctgcctatga atctgtgaaa tcgaaagcaa gttagttact tcctagatac aataggggta  104460
catgcattgg gtaaatacac ccattccaaa tgggagacat tggccaaaat aaaggggcta  104520
caggccccat gcaagtctga aatccaatag ggcagtcatt aaaccttaaa gttccaaaac  104580
gatcttcttt gattccatgt ctcacattca gggcacattg atgtaagagg tgttctccca  104640
tggccttggg aagctctgcc cctgtggctt tgcagggtac aaccccccct ctggctgctt  104700
tcatgggctg gcattgagtg tctgcaggtt ttccagatgc acagtgtagg ctgtcagtgg  104760
atctaccatt ctggggtctg gaggacggta gccctcttct catagctcca ctaggcagtg  104820
ccccactggg gactctgtgt aggggctcca gtcccacatt tcccttccac actgccttag  104880
cagaggttct tcatgaggtt cctgcccctg cagcaaactt ctgcctggac atccaggcat  104940
ttccatacat cctctgaaat ctaggtggag gttcccaacc tcaattcttg atttctgtgc  105000
acccgcagac tcaacaccat gtggaagctg ccaaggcttg gggcttgcac cctctgaagc  105060
catggcctga gctgtacctt agccccttttt agccatggct ggtgcagcta ggatacagta  105120
cactaagtcc ctaagctgca cacagcaggg gggccacaca gcagggggac tctgggcctg  105180
gcccatgaaa tcatttttc cttctgggcc tccaggccta tgatgggaag ggctacgttg  105240
aaggtctctg acatgccttg gggacgtttt caccattgtc ttggcaatta acattgagct  105300
tcttgtcact tatgcaaatt ctgcaaatgt ctgcagctgg cttgaacttc tccccagaaa  105360
atgggttttt ctttttcttc ttccttttt attattatac tttaagttat ggaatacatg  105420
tgcagaacgt gcaggtttgt tacataggca tacatatgcc atggtagttt gctgtatcca  105480
tcaacctgtc atctacatta ggtatttctc ctaatgctat ccctcctgta gcccccgact  105540
cccctgacca gccctggaat gtgatgttcc cctccctgtg tccatgtgtt ctgattgttc  105600
aattcccact tatgagtgag agcatgtggt gtttggtttt ctaatcctgt tttagtttcc  105660
tgagaatgat ggtttccagc ttcatccatg tccctgcaaa ggacatgaac tcatcctttt  105720
ttatggctgc atagtattcc atggtgtata tgtgccactt tttctttatc cagtctatca  105780
ttgatgggca tttgggttgg ttccaagtct ttgctattgt gaatagtgct gcaataaaca  105840
tatgtgtgca tgtgtcttta taatagaatg atttataatc ctttgggtat ataaccagta  105900
atgggattgc tgggtcaaat ggtatttctg gttcaagttc cttgaggaat tgccacacta  105960
tcttccacaa tggttgaact aatttacact cccaccaaca gtgtaaaagc gtttttattt  106020
ctccacatcc tctccagcat ctgttgtttc ctgacatttt aatgattgcc attctaactg  106080
gcgtgagata gtatctcatt gtggttttga tttgcatttc tctaatgacc agtgatgatc  106140
ttttttttta tatatatttg ttggctgcaa aaatgtcttc ttttgaaaag tgttcatatc  106200
cttcacctac tttttgatgg ggttgtttgt ttttttcttg taaatttgtt aaagttcctt  106260
gtagattctg gatattaacc ctttgtcaga tggatagatt gcaaaaattt tctcccattc  106320
tgtaggttgc ctgttcactc tgatggtagt tccttttgct gtgcagaagc gctttagttt  106380
aattagatcc aatttgtcaa ttttggcttt tgttgccatt gcttttggtg ttgtagtcat  106440
gaagtctttg cccatgccta tgtcctgaat ggtattgcct aggttttctt ctaaggtttt  106500
tatggttctg ttgcatcgtc aggctgcaaa ttttccaaac ttttatgctc tgcttcctct  106560
tgaacacttt gctgcttaga aatttttttcc accggatacc ccaaatcatc ctctcaagtt  106620
caaagttcca cagatctcta gggcagggac aaaatgccac cagtctcttt gtatagcaag  106680
tgtgaccttt actccagttc ccaacaaatt cctcatctcc atctgagacc acctcagaag  106740
ccattcaaca ggtctctagg aagttccaaa ctttcctaca tcttcctttc ttctgagccc  106800
tccaagtctc taggaagttt caaactttcc cacattttcc tatcttcttc tgagccctcc  106860
agactgttcc aacctctgcc tgttacccag ttccaaagtt gcttccacat ttttgggtac  106920
ctttacagca gcaccccact acccagtacc aatttactgt attagtctgt tctcatgctg  106980
ctaataaaga catacccgaa actgggcaat ttataaagaa aagaggttta atggactcac  107040
agttccatat ggctggggag gcctcacaat catggtggaa ggtggaggag gagcaaaacc  107100
atgtcttaca cggcaggagg caagagtgtg catgcagagg aactgccctt tatgaaacca  107160
tcagatctca tgagacttat tcactatcat gagaacagca caggaaaaac ccacccccat  107220
gattcaatta agtcccacca ggtccctcct aggacacatg gggattatgg gagttacaat  107280
tcaagatgag atttggatgg gaacacagcc aagccatatt atatacttta ttaactactt  107340
atctcttcta ataatacact tgttctcatg tggttattct tcttattacc tttttttga   107400
gatgaagtct cactatgttg cctaggctag tctcaaactc ctgggcatga gccatcctcc  107460
cacatggcct cccaaagtgc tgggattaca gatgtgagcc atcactcctg gtctcatgtg  107520
gttattacag tgaatgatta agacaccttc tctccttcac tttgtcccta ctttttctca  107580
```

```
gatttttttc aatgatgtct atgcttttct tgtttttttc tagcttttct aaccttgcat  107640
ttatttttcct tcagatctca acatctgctg caatttgact gttcaggtta gtgacaattt  107700
gcccatttat cagttttgtg ccttaggcag tattcaccag cattctccta cttgagatga  107760
ataaggatct ttatttatct gaccacttgt ttactcattc atggggacat ttaatattta  107820
cagaacactt tcatcaaaac aagcctgttt tttcttttca aaatataata tactagcata  107880
ggaacttgac agaagaggta ataatacaga agaaatctag agaactgatc atggagaaat  107940
aattaaacta aaacaaagct gctgcttata gtaaggtaga ccaagtttgt cctgtgttcc  108000
aaattatact tagccaaaaa taaatattta tagataattg aatagtagtt tttagaaatg  108060
attcatggat tactcagggg tggaaattat ccctgtaatg taggccccaa acttctaaaa  108120
tatttataat ttgtgaggga gaaataaatc cacaaacatt tgaaaatatc taattttaac  108180
ttaaattgaa aaacagcagt agtgtatttt tgttaggccc ttatattggg taatatagtc  108240
tatgtaagta tgggaaaggc tggtgaaaac tgtggattta tctacaaaat acattggtac  108300
attggggagt tttgtgtggg aagtgtccta acacctcagg aaatgctaag gaagaaatgg  108360
gtagggtaga ctgaacactc cacaagggta gggagagcgt cttctcttca ttgctctagc  108420
cccagcacct agaatgtgcc tggcactcaa tcttttcttt ccttttttaa ctgatttcag  108480
gttgagttta taggctagct gagtttatag gctggtagag tacaaggtct actaatttct  108540
ctactcatga ctactttact ttctaatagc actttatgtt cttcaagtgc ttttttttctt  108600
tgttcttcaa tgtataataa ggtgagaaca gttttatctc catcttgcag ggagaagagt  108660
tagttaagag actttctcag gtcacacaca tatagttaat gacaaggtga ggtttaaacc  108720
ttaaataata gaaataaaag tgattttata attatctaga gtagtttcaa tgtgaaataa  108780
cttaaaggta tggaaatgga tgccaagaag tatagtcagt cttgctggag taaaataatg  108840
cccagtgctt tgtgccttct cccagctgct gcttccagaa gaacggggtg tctgagtgtg  108900
aacatcaccc aacaagtagg ttaacagata tccacgcccc tcttgaccca catacatatc  108960
agtgggattt agaatgctgc cacatattga tgattgaatt tatgaagcat ataatatcct  109020
caataataaa ccaagtgtcc ctgtcccaac ttgttatctc tgcttctgtg aacacatgtt  109080
ttcttttata tgctccttac tcctcaggtg ctctctcagg gacttttcag ttcctgacct  109140
tgtccttttc agcattttct cagaggacaa ttcttagctt cctgttgatt cctcaagcat  109200
taattgcttt tttctgccag atatttcctt gctaggctct tgagccctca gagctgttct  109260
gaattatgca gtgggaattg cccaggatta ggaatcacct aatgtcccca ccacccccgc  109320
ttccttgtga ggcacttctc aactctgcat cccttatacc ttcacagcaa ccctgtgtac  109380
ccaaagcagt gtcatactcg gtgcctcctt ttctctcctg aaataaaatt cctagataag  109440
aagacctcta tattccaggc cttgtctttg attttaggga aaaaaaagaa aactacctat  109500
atacataatg tttttaaaaa tcagtaatgt cccactcgtt acagaaagga gaaataaaag  109560
aagtaagtta atgcctggga tacgtgctac aacatggatg aaccatgagg acattacacc  109620
aagtgaaata ctccaggcac aaaagcacga atactgtatg gttccgctta gatgaggtac  109680
ccagagaagt cacattcata aatactgaaa gttgtatggt ggtttccaag gggaggggga  109740
aatgaggagt tatttaatgg gcacagagtt tcagcttgag aggaggtggt gacagttgta  109800
caacaatgta aatgtactta atatagtaca cttaaaatgg ttaaaatggc aaatttttatg  109860
aaataggaat ttatcacgat aaaaaattaa aaagtaagaa aagttactgc ttgggcgaaa  109920
gtatatcaaa aaaataaaaa tagtccccac aaatttccaa aacaacccta atgaggtgtt  109980
gctgcctaaa tggtgaacca aattgtgaac caatgtgtag tgtttgagac tgggaaactg  110040
atgcccaaga ttttagcctc aataaggagt agagttcata atttgactcc aaagacattt  110100
ctttccctac catgccaagg ccatctgatt cccagtccaa agaagttttc tctctgctct  110160
gtaggctgcc ttaatccaga gtacacaagc cttccatttt cttatctgtc ctcctaccag  110220
ggtgtggtcc ttttcctcct gaacactgac tgtataatta ccagacaaaa ctaaacatat  110280
ttaaaatata ggcagtcctc tacatccaag gttccacatc cttggattca accaaccatg  110340
gattgaaaat atttggggga aaaaaaaaca ataaaaaaac actggcctgg gcagcatagt  110400
gagatgccat ctctacaaaa acattaaaat attagctgag cattccagca ctttgggagg  110460
ctgaggcagg cagatcacct gaggtcagga gttcgagacc agactggcca acatggcgaa  110520
accctgtctg tactaaaaat acaaaaatta gccaggcatg atggcagctg cctatagtcc  110580
cagctactca ggaggctgag gcagggaaaa tttcttgaac ccgtgaagca gaggttgcag  110640
ttagccaaga tcccaccact gcactgcagc ctgggtgaca gagtgaaact ctgtctcaaa  110700
aaaataaaaa taaaaataaa taaaaattag ctgagcatag tggcatgtgc ccatggtccc  110760
agctacttag ggggttgagg tggcagtgag ctgtgatcgt gccactgcac tccagcctag  110820
gcaacagcga gaccccatct caaaacaaaa acaataaaac agaacacaga ttaaaaacaa  110880
aatacaggct gggctcactg gcttatgcct gtaatcccag aactttgaga ggccaaggtg  110940
ggaggattgc ttgtgctcag gagttttaga tcagcctggg taacacggca agaccacatc  111000
tctacaaaca acaacaacag gagactatac tttcagggac catttctggg gatcatagtt  111060
tgttactaga gaagtttctc tgtgtagagc attgaaatat aaaaatgcag aataatcatt  111120
tacatagcat ttacattgta ttggttatta taagtaatct agagattaat taaagtatac  111180
```

aggaggatat acataggtta catgcaaata ctacaccatt ttatataggg gacttgatca 111240
tccatagata cgggtatctg aggaggtgtt gggttcagtt ctccacggat accaagagac 111300
taatgttaat ttcatttccc caacctccac accagaactc tgaaataaga ataagaaaag 111360
gagcagttgg gatagacaat atcagaagta tgtggaaatg ataacagtgg aaggaaagct 111420
gatctaggcc tactcaacaa attttaatct tcattctggt aaaaacaaat tagatttatg 111480
ggtgcaaatt tgagccagca attagatggc tcttaggatt aataaaaaaa gactgaacat 111540
catgccttcc aaagactgag ggaaagagat agataggaga ctttggcaaa gtagcacttt 111600
agccaacatc attagcctaa atcttagtga agagaggtta gaagaaaggt agaattttca 111660
tggaaggatc catttttctt cacttcagaa ttaagggaaa aattaggaag ctgaataaga 111720
actaatggcc taatttcttt gtttctttca aaaatcaaat ctttaagtta aaatttcaat 111780
aaccaacaga agaaggtaga accaattttg tgaatttaca aaatacttgc ttattgacac 111840
tattgccaag gtcattaaat aagttatcaa gttacctcag atcccagtga tttaaataat 111900
gctttctgaa tgtatccttt tctgttttaa gaagaagctg tattaggttc ttgtgttcct 111960
ataaagaaat acctgaggcc ggatgattta taaagaaaag aggtttaatt ggctcacgat 112020
tctgcaggct gtataggaag catggcccca gcatctgctc agttttctgg tgaggtctcg 112080
gggagctttt actcatggcg aaagcagagt ggaagcagca ggtcacttga tgaaattgag 112140
agcaagagta tgggtgggga gctgccatac tcttaaccca atctctagtg aacacaagca 112200
ataactcact tatcaccaag ggaatggtgc taagccactt gtgatggatc cacctccaaa 112260
atccagtcac ctcccaccag gtcccacctc caacattggg aatcacattt caacatgaga 112320
tatggaaggg acaaacattc aaaccatatc agaagcctat cttaggctgg gcacggtggc 112380
tcacgcctgt aatcgcagca ctttgagagg ccgaggcagg cagatcattt gaggtcagga 112440
gtttgagacc agcacgggca acatggtgaa accccatctc tactaaaaat acaaaaacta 112500
gctgggcatg gtggcacaca cctgtaatct cagctactcg ggaggctgag gcaggaagct 112560
ctcttgaacc cgtgggcaga ggttgcagtg aactgagatt ctgccactgc actccagtct 112620
gggcaaccga gtgaggctct gtctaaaaaa aagagaagcc tatattaaac ttataaaatt 112680
taatatcatt tcaactagcc ttttgttggg tgcatttgtt cactttggac tatttttcca 112740
aattcatgta cagtcgtgca tctcttaaca atgaggatat gttctgagaa atgcatcctt 112800
aggcaatgtc attgttgtgc aaacatcata gagtgtactt agacaaccct acatggtgta 112860
gtcactacat acctaggcta tatggcatag gtagagccta ttgctcctag gctacaaacc 112920
tgtacagcat gttactgcac tgaatgctgt aggcagttgt aacacatggt atttgtgtat 112980
ctaaacatgg aaaaggtaca gtgaaagtac aggattataa tcttatggaa tcgctgttat 113040
atatgtggct catctttgac cagaaatgtt attatgtagc acatgactgt aattccacta 113100
ttgattagag actccataac ccacacctac ctgttcattt gcatgcacat ttagccgata 113160
ggtgacttta tcactagatc agtccaagaa taagtttaaa gagcatgtcc ccttttcttg 113220
ttttttaata aatgtaaaac atggaagtac atgatacctc aaaggagcat gaatcacaaa 113280
caggagtcca catttacaga gcctttgcaa ccacttttat atcttcaaag tgctctacca 113340
aagtatcgga gtcagtcagt catgcaggta ggctgccaac catggctgag ccctcaggag 113400
ctgtcttgat gaccagcata gacacggtct gttgtgttca ttcaaacttg ccaggaacca 113460
aggatctgag aatggtaagt ctggtgtgcc aagaatgaaa ctccaagaaa gaaacttcag 113520
aatcaaattt aaatttaggc tgggtgtggg ggctcatgcc agtagtccca gcgctttggg 113580
aggctgaggt gggcagatca cctgaggtca ggagtttgag accagcctgg gcaacatggt 113640
aaaactctgt ctctgctaaa aacacaaaat tagctgggca tggtggtgga cacctgtaat 113700
cccagctact tgggaggctg acgcaggaga attgcttgaa tccgggaggt ggaggtggca 113760
gtgagccgag atcacaccat tgcattccag cctgggtgac aagaccaaaa tgccatctca 113820
aaaaaaaaaa aaaatcgaat gtaggtggaa ttaataaaaa tttaaataaa ctaacagatt 113880
aaacctttca ttgttcagga aggaacagaa caatttttga taacttgtga aatatctggc 113940
acagaaatta tttagagcca ctaaataatt tcaaattacc taaaaatcct agtgatttat 114000
tttctatttt aagatgaagt ctactttaaa cttctaaaat gcagggttat ttaaactggc 114060
atctaaatcc aagctggttt tggttggtaa ttcctctagg acattttact aaatcttgat 114120
cttatctaaa tgatgctatg tcatagatgg actgtttgtt ttgtttgtaa tccagggaaa 114180
ttaaaaaaaa aaaaaaacaa gtagaaataa aggcttttaa agaattttta gagttagaaa 114240
tgttttcaaa attaggttct ttaaaccatt agccatctct tccttctgaa ctcttctttt 114300
ttctgccctt tggtagctat gaaataatct gcattccaga aacttctttt ttcccagtcc 114360
tttttcatgt cttaacagtg ccatgcatga ttatctacac catggaaacc catcttaatg 114420
aaatggaaag atctgctgtt taaaaaaaca aacaaccaat cacccatgcc tcctacagtc 114480
ctgttaggcc aaaggtctgc tctccacctg gtcagctgct gaggaggggg caggtacctg 114540
taccacagtt tagccacaac agaaatatcg ttgggctacc aaaagatgtc agtccttaaa 114600
aatgttggca agcaaagaaa tttttctcta agcacatctt agagatactt tagtctaaag 114660
gcacagttga tcagattaaa gaagctaaga ttacattttg gggctggtct aatcgataaa 114720
tatctatcag tggccataaa gttttatttc agtgttttct ggtgtaaagg gattaatttt 114780

```
atgtactcaa gtctatggag tttgtttttct tccaccttta atcacctcac tcggcttgct  114840
ggtgtgcagt gctcatgcaa aacttcctaa gggagtgccc cacttgtttt aatttacttc  114900
ctctagcaaa aacacagcta gaggaaacac aaatcgttta aaaaagaaag aaaaaaacaa  114960
aataaactgt ctcatttcat tcagaagatt tctgttcatc ctacttttaa gtcaaagggt  115020
agtgttaagt tttgatttgc aacaaactca gattatccat taacccacgc ttgcaggctt  115080
tctactccct ttcagatgta cgtctatttt tttaaagggc ccaggctttt tgagtttacc  115140
tttcaaaaaa gtatgtttag gcagaatggt cttctagagt tatggaatag cttttcaaat  115200
gttggatgga tattttctta gaattactcc aaaatgcagt aaatgagaag caacaagctg  115260
aagctgccag tctctggtat gaatagcagt gagtaaaagt gctttgaact tagtgtccgt  115320
attggaaagc aggaagttga aagaatcagt ttgaattcca taattgctgt aaatattatg  115380
tagcatttat tatcccattt tttgatgtac catgtaaatg taaaacatac aggtcccctc  115440
cctcccctga agagtttatt ctcctagaca taaaatgagt gtacatgcct agccattgtc  115500
caccccttc cagttcgatg atgtgtatct gatgaggaag gacagagatg agagaatcaa  115560
aatcattata aacgaataca gatgatgact gggtgactga aatcttatct cccaaagagc  115620
atagttaact gcagcagcag tggatgataa ctactgggtg ggggtagagg ggttgttttg  115680
ctcagttctg cctagaagac agttgtagtt attttagtcc cacagtctct actcctccct  115740
gggcctgttt ttgccctgct ttcctggcct ctgtatacgg cctttgttag cactttgtag  115800
ttgtcagtca gctgttactc caagtcctct gtcagaaata aagttacttt tgtttgaaga  115860
gtctccagta atccccctc ttttaaact atgactcccc aaagatatat agtctaactt  115920
gtgtgcacca gtattttatc atatttattt aataaattcc ataaacctta cataaaagat  115980
attaaacaaa aagtactcac ccccttaaaa ggaggaagat aatgaccatc aaggcatggc  116040
caaattacac agtcttagga aacagtacag taggtaatta ttggaccttt ttgattaatg  116100
gcatccttgc tttgcaggtg tgtaccacca ctgcacttca caggtttctg taaataaagt  116160
gatttggggt tcccttctgt tgagcattct tctatcgtga aaccatttgg tgatgagaag  116220
cttgagtttc taatgcatgt tcgttggctt tattggagct gcttttgtac tgtggccacc  116280
catttaactt cctctgtggt tgatatatga gagctagagg actctgcctt agcttcttca  116340
tcagaacacc atgtgatgtc tccaaatgtc ttacactgtg gctcccctgt aagagaggtg  116400
gagaggagag aggccaaggc tgaggccatt ccagtgagat gatcactttc ttaggctact  116460
gcctgagacc tctcaaacag gatctgtctg tgggctctag tgaggggaaa gtaaccggcc  116520
ttcagttggc tggtcagaaa cagttaccca cggccgggtg tggtggctca tgcctgtaat  116580
tccagcactt tgggaggtca agggtggagg attgcttaag cccagcagtt tgagaccagc  116640
ctgggcaaca taacaggacc atgtcactac aaaaatgatt taaaaattag ctggatgtgg  116700
tggcatgtac ctgtggtccc aggtactcgg gaggctgagg taggaggatt gtttaagccc  116760
aggaggtcga ggctgcagtg agctgtgatt gcgccagtgc actcaagcct gggcaacaga  116820
gtgaaactct gtctcaagaa gaaaaaatgt ttcaggcaca gtcagtgttg aagatgtcaa  116880
ttagcaaggt ttttttaatc ccctggaagt gtgcaaccca ggattaaaca gatgtttaag  116940
aacataaata aaagatgaat tcctggccga gcgcggtggc tcacacctgt aatcccagca  117000
ccttgggagg ctgaggcaag tggatcatga ggtcaggagt tcgagaccag cctgaccaac  117060
atggtgaaac cccgtctcta ctaaagatac aaaaataagc cgggcgtggt ggcgggcacc  117120
tgtaattcca gctactcagg aggctgaggc aggggaattg cttgaacccg ggaggcggag  117180
gttgcagtga gctgagatca tgccactgca ctccagcctg ggtgacagag tgagactcca  117240
tcaaaaaaaa aaaaaaaaaa agatgaattc ctgacactta atacatttag ataaagtcta  117300
acttatcttt aaagttaact ttaaacaact gctctacagc ataatattat cccctgtgta  117360
gacagtgacc tctagtataa gctttaaatt ggtttaaaaa atagattaag tatgaaaaca  117420
tctcttcaac attttcatat ctctacaact ctttattaga aatatccatt ttgtcctata  117480
ttcactttct gctgaactaa agtgcatgcc agtctccaga ggatgacaga tgggcagaaa  117540
tttataaaac agcagcagtc atggacagcc taggccacaa gtagaacata ccagacctcc  117600
tgggtactac tgtctcagtg agaaaggatc cagagatcca gcatcaccag attccctcac  117660
atactgacca aagaaacttt tttagtttgc agattttggt gggaataggg aggagtatta  117720
gaagctgaag ttcttgcttt atttaaataa tcatggcagt aacaaaattc ggttaaataa  117780
tgccccttct aggccaggcg tggtggctca tgcactttag gaagccgaag caggcagatc  117840
acttgaggtc aggagttttg agaccagtct ggccaacttg gtgaaacttc atctctacta  117900
aaaatacaaa aattagctgg gcgtggtggt gcacacctgt aatcccagcc actctggagg  117960
ctgagggagg agaatcgctt gaaccaggta ggtggaggtt gcagtgagct gagatcacac  118020
cactgcactc cagcctgggc gacagcgaaa ccccatctca aataaataaa taaataaata  118080
aataaataaa taaaataatg ccccttctag ttgaaaaact aagttcctac ctaagcataa  118140
tttggattta cccaatttat cttctttcaa aatacctcaa acatttcacc ttattattct  118200
ttttaaggat tacaaagtag agcaggggggg aaataataaa ccactaataa agaataatag  118260
ccatttgaca gacaggtgtt cttagtttca taaaaaaaaa gatgcctggt agattcagtc  118320
ttttatgaat actaaagaat gcctcttatt tttgtttgtt ggagacaggt tcattttgaa  118380
```

cctaacctgg tgttctcagg gaacataggt tagaggggag agatttagga gtgagggctc 118440
aagcaagaag catgttagaa gactgctgta gtggtcccgg tgaggagtga aaggaatgga 118500
actaaaatac catgaggaac gtgggccaaa ggaaacaagt ctgagagatt tagaaagtaa 118560
atcatcagga ttcagtggtg actcttggat ttgaagaggg aaagggaata atctagggta 118620
gctgtcagtt tctgccatgg gtagttgggc taacagtagt gtattaactg aaatgggggg 118680
caggcaattt gtgagggtga gttgagttca gcacagggcc tgttgacttt gaggtgcctt 118740
tgaaacagga gtggatatgt ctaattttac atgaagtgtc tacttaagag aatgtgtaga 118800
gacattaaca gggctgggta ggaacacgga atacctcagg tgccaaatga aaacttttgg 118860
caataacaag aagttaggaa gttgaagagg gtagaagagg cagcagaaat gtagatggat 118920
aagaacatgt tggcatgttt aagtgacatt gtgaataaac ctgtcctact gggccagatg 118980
ggattagggc tgagatcaaa gcaggcctaa tgcagtttgc agacgtgttt tgttagactt 119040
ttgtagagct ggatcacaca gtgtcctaaa tttaaattag gtgccaacat ttccacacaa 119100
aatccggatt tctgactttt cttttaaact aaagggctcc tagaggtaga tttggcaaca 119160
ttggtagacc tatatgataa taatcgactg aagtatatgt cctttctctc caatgaacct 119220
gttttactta tgttatttcg ctgagccctt agacatttaa attttgtact ttttttttt 119280
aatccaggct tataagtcag atgaattttc tacttctgag tcaaagatca gtaggtaata 119340
aaggtacaaa gatagattag caacagatta cggagagctt tgaataccaa tctaaggagt 119400
gtaaatgtag gcagtgggcc acctttgaat aaggaattga tgagattaaa gccatattta 119460
ggaggattat tctggaccag tatgaaaaca cagaagttag ggaaaacagt taatagtttt 119520
gaaagagaag agaaaaagga gatggtgttg ggatacataa atgggctttt aaaatgcaaa 119580
atgagaagtg ttttaaagag atatcaccca gaaagtctat gcactgccac atgggcacta 119640
tatgggtggt tgtatttggt gggaaatttg cttgcagact tccagaactc agaccaatgt 119700
gtggtgtggg ggacggtgat tgtcaggcat tatggaaagg tcaaacaaaa tatgctcact 119760
ggctatctat ggcccacagg tcactgtagt ctctgttata agtacactaa gtggaggaga 119820
aaggtccttt aaaaaaaaga aagctaaaat taatacctga ttgttattaa ctgtgtgcca 119880
aacactgttc taagctcttt acacagacat tttatttaat cctcgcaacc aatttctgaa 119940
gtaggtactt ttccaatttc cattttacag acaaagaaac tgaaacccta gaggttaaga 120000
agttatccaa agccacaagg ctgataagaa cagaaccagg acttgaacgc aagcagtctg 120060
cctctccaga ggtttatctt ttaactgcta tgttaaactg cccctgcatt ttaatctgtt 120120
ctaatgctac acagataggc aactttacag gtagaggacc ttatgcttta ttctggatgc 120180
tctgttataa ctcgtttcag gggtgtcaat ttggtccagg tcctcctgga agaaataaaa 120240
ctccagaatt gacccttgaa ctgtctctta gggagcagat aatgtaacgg gtccttgggt 120300
gaccttgaga gaacaggtat gttcaaatgt ctgttctctt cctttagcta atggatcagt 120360
gtagcttata attgcatgct tctaaccctt tgttgaaaaa taaaaactct tataaacatg 120420
ctttttttt tttttggaga cggagtcttg ctctgtcgcc caggctggag tgcagttgtg 120480
caatcagctc actgcaacct ctgcctcccg ggttcaagct attctcctgc ctcagcctcc 120540
cgagtagctg ggattacagg catgcgtcac cacgcctggc taatttttat atttttagta 120600
gagatgaggt ttcaccacgt tggccaggct ggtctcgaac tcctggcctc aggtgatcca 120660
cccacctcga cctcccaaaa tgctgggatt acaggcgtga gccaccatgc ccggccttaa 120720
aaatgctttt aaaaatgaaa actaaaacat gttaattttt ttcaaatgtt ttcatgaaaa 120780
ttatcacagg acaagtttca taaatattga aatttggaaa aagttgcaag cctataacat 120840
tgcagagaag caaatgcatt tgatgcaaag cctcaaattt gtcaagtttt tctaccatat 120900
tcagtgtggt ttctttctct ttggcctata gatgaaacat gtaaatgaaa gatttcaaga 120960
tgaaaaaaat aaagaggttg ttctcatgtg cattggcgtc acttcaggag ttggacgact 121020
gctctttggc cggattgcag attatgtgcc tggtgtgaag aaggtttatc tacaggtact 121080
tttttacacc ttttttcccc tatcaaaaat tactctcatc acccaatgtc tcattaaatg 121140
tacttacatg cttaaattct ttttttttct ttcctttttc tttttgagat ggagtttcgc 121200
tcttattgcc caggctggag tgcaatggca cgatctcagc tctccgcaac ctccacctcc 121260
cgggttcaag ggattctcct gccttagcct cccaggtagc tgggattaca ggcgtgtgcc 121320
accacaccag gctaattttt gtatttttt agtagagatg gggtttctcc atgttggtca 121380
ggctggtctt gaactcctga cctcaggtga tctgcccgcc tcggcctccc aaagtggctt 121440
aaattcttct ataaaaatga gaaatatttt ctacaacata acttctatag gcagtttttc 121500
aaggacaaaa ttagttatta gtttgggttt taaacatgag aaattggcaa tgaaacaaca 121560
atttctttgt tttgtcgtgg aactccacca aaccagaatg gttttcatcc attgcttttt 121620
ctatgaagaa tgtttttttgg tgtagttctc atagtcatgt gcagatcctg tgccctttgc 121680
atgtcttatg aaatttggtt gtgtgtgtga cttttcagct tctttactgc aaattgcctc 121740
ctcgtgtttt ggggtgagca taaacaaatg ctaattccaa gatcattgct gacaatcaac 121800
agaacaggta ttgaagtgac tccttcatgc cacacactct gctaaatgct gaagacttaa 121860
gtgaaacatg gtctgtgccc tcacctagta gctaatggtc tcatgggaga agataaagca 121920
gtgttgcagc acagtggaat aacgggtttg ctagatgagt gccatagcaa cacaggcacc 121980

atgcatctga gtcactagtg aagggctgag cagagttaag aggtgaggac caagtgtgta 122040
gataagggag acagagaagc atgctccgag cagagacaaa cacatgggaa acacctccca 122100
cactgtggag gtgtgacatg ggatgttata tctgggaaac aaatgtttga atggaataaa 122160
gggagaatag tggatggatt ggttgggggg atggacagga agcagcttgg gaagggcggt 122220
tatgtacatg agctctatcc tgcaatctac taggagctat taaaggattt taagccagag 122280
agtgacataa attggggtgg cgggggttgg tgtttttgg ttttttgagg caaggtgtca 122340
ctccgtttcc caggctggag tacagtggtg ccatcacagc tcactgcagc ctcaacatcc 122400
cgggctcaag caatcctccc acctcaaact cctgagtaga tgggactaca ggtgcgtgcc 122460
accatgcctg gctaattttt gttttttgt ttttttctg tagtaacagg gtttcgctgt 122520
gttgcccagg ctggtctcga actcctggac tccaacgatc cacccatttc agcctcccaa 122580
agtgctgggt ttacaggtgt gagccatgcc cagcctgagt tgtgttttat aaagatcaat 122640
ttcggctgtg cgtggtggct catgcctgta atcccagcac tttgggaggc tgaggcaggt 122700
agatcacctg aggccaagag tttgagatta gacatggtga aaccccgtct ctactaaaaa 122760
tacaaaaatt atccaggcat ggtggcacat tcctgtaatc ccagctactc agcaggctga 122820
agcaggagaa tcacttgaac ctgagaggca gaggttacag tgagctgaga ttgtgccact 122880
gcactccagc ctgggtgaca gagcaagact ccatctcaga gaaaaaaaaa aaatcaatct 122940
ggtaactgtg aagagtggaa tggagtggag gaggggacaa gagttgagat agtcacaaaa 123000
ggccacatac tgaatgattc catttatatg aagtgtccag aataggaatc tctgtctctt 123060
tccatagaga aagaaggtag atttgtggtt gtaggggcta gggaaaggag ggaatgggga 123120
gtaactgcta ctgtgtttct ttgtggggta atgaaaatgt tataaaatta gatagtgatg 123180
aaactcactg aattgtgcac tttaagtggg tgaatttcat agaatatgaa ttatatctca 123240
ataaagctat ttttaaaaag gccaggcatg gcagctcatg cccataatat caacacttta 123300
ggaggctgaa gcaggaggat cacttgaggc cagttcgaga caatcctgga accatagtga 123360
gacctcatct ctaccaaaaa aaaaaaattt ttttaagtta gccaggcacg gtggcacatg 123420
cctatagtct tagctatgca ggaggctgag gtgggaggac tttgagccta ggactttgag 123480
attacagggg agtatggtca tgccactgta ctccagccta ggtgacagag tgacaccatg 123540
tctctaaaat ttagagaaaa agagtagagg tccagggact agttggaaac tatattaaag 123600
tagtctaggc atctaggcag gaagaaccaa ggcagagata ctgagatcaa atggaataag 123660
ctagggacat agagaatgag aatttgatga agccacttag gctggatctg agatttgtgg 123720
ctcaggtaat ggagtagatg attttgccat cttatcaaga tcaatatgca gtagatctaa 123780
gtatgatgat gaattggggt tttttaaatt aaaaaaaaat tttcacccaa cctacctgcc 123840
cttgccccag tgatgagttt gtttttgcca tgttgtttgt gatgcctgta ggaaagccaa 123900
gtagtgttta gagctcatgg aagagtctgg ataagaggcc tgtgggttaa gaggcctgtg 123960
ggttaggaag cctgtgggtt tggcattcat gggctgcatg aagtaagtgg aggactaagg 124020
agatggaaag ggagtggcca gataggcagg gggaaaacag ggatgagagt gtttctaggg 124080
gcaggagtgg tcagcagtgt cagtagctag tgacaaagag gctgaaaggt tttcattgaa 124140
tttagaatag ggaggctatg agtgagctta cagagaatgg tttctacagg tgcagaaaat 124200
tgatcatgat catttgaaca aatggaaata tagacaactt tgtccaaatg cttggctgtg 124260
gaaggaaggt aaaggcaagc cacaaggggg ggtcttaggt tttgaggatt aagcctgtct 124320
aattatatta tacatgagag gcagtatttg agcagggcct tgaaaaagag gtgaaatttg 124380
gacacaggaa aatgattgga aaaggcatta cagataaagt taacttccat taattgactt 124440
gaagtaataa cagtcaacca tttattgagg actttcatgt gccagacaat gaactaaggg 124500
ctctacatgc attatctcat tgatccttgc cccagccctt taagagagaa ggtaccattg 124560
ttatttccac ttagagatgt gaacactgag gaactgagag gctacctcac tgtgggtgtc 124620
tgtgtgtata gaggtccagg cagtcacagg acagtctggt cacacagcta gaaggaagtg 124680
aagagagttg gtaatgtgct gcttttaaaa atgtatttat tgtatcatga tcactttgtc 124740
agtacttcac tgtggacatc ctcatctaac atttagtttt gttctctagt gtcaagggag 124800
ccctctaatg gacatttatt gcactacaga ctttcagctt tcatacattc aaaaattgag 124860
tgcctcctgt gccccaagca ccagctcaga tgctgttagg gtgatgcaaa caagacagac 124920
atggtccctg agttcctaaa gcaagcctga ggcaggaaga agccgaatgt gtgtggaaac 124980
ccaagaagat ggggaaagtg gcatgggaag ggactggaaa gttagagtgg gtcagattag 125040
acagagcctt gaaagccagg atgaagagac tttgctttaa gagtagtgga ttttggccgg 125100
gcgcagtggt tcacgcctgt aaccccagca ctttgagagg ccaaggctgg cggatcacga 125160
ggtcaggaga tcgaaacaca gtgaaacccc atctctacta taattacaaa aaattagcca 125220
ggcttggtgg cacgcgtctg tagtcccagc tactggggag gctgaggcag gagagtcgct 125280
tgaatctggg aggcagaggt tgcagtgagc cgagatcgca ctactgcact ccagcctggg 125340
tgacagagtg agactccgtc tcaaaaaaaa aaaaaaaagt agtggacctt tgttggagga 125400
cgtaggtctg attttgtttt taaagtcact gtggttgttc tgtgtttgga ggatggatta 125460
cagggaggga aacaagagtg caagagggga ggaagacctg ttaggagatc aattcagggt 125520
ccaggtgaga gatgatggac tatggtgctg taatagataa agtcatgata tattttatat 125580

atcagcattt tattccaaaa caatgtttga acgtcctcct accctagata gggcagacag 125640
attatctgca acatttttgg agcacatttt aatacctgac tgtttccagt aatttacaaa 125700
agaaaatata gcctttccta acttgtccca tgttggtctg cagttacaca gcagtaagtt 125760
aaaagttagt attgggggtc aaatatttca ctttagatga aagtttagcc acaatctggc 125820
ttctgttagg ccttatctaa tttttgcatc caaatgtaga gcatcgtttg tggcacccag 125880
tagcacatgc tgagtcacag gtgtgacagc tgcatttcaa acaagcctga gaaggagaaa 125940
gaaagccctt cagtgtgtcc tgtggttgcg aggagccact cacggactcc accttgtgaa 126000
cacagcggca caggacgcaa cacaggccta acccatgcag gatgctggac tcgttccctt 126060
attcactacc tcctctctct ccttttttcat ggcttccttg ccccaacatc ccaaacacac 126120
agtgtggttt ttgattcttg gctcttctct gctgcagttg actccagctc tgcctgtttg 126180
ttccttcctt tttctccatc cctggctctc tgccttttgg cccatccttt aaggcttgga 126240
atgctcctgg gcttaccttc tttccattca ttggtggtgt gttttctcag acatacctgc 126300
tccctgcttt catctttcaa cttcttgggg ctgtgacaac tcttccctct ttgtcccctg 126360
gaagccagtt ctgagtagca gccaggccta gaacactggt gacacagaca cacttcatag 126420
ccctcccgca tggtctagtt tcagatcatg gtaatcccta gtctaggagg ctgccgagcc 126480
caagagcaca ggctctggag tgagaagcca gttcacccca gttctaccac caacttgaga 126540
atcagcagag ggctgtggtg aggattcttg gtggcaggtt atgtaaagtg cctagcccag 126600
actggatgat taataaatcc ttgcatctgt tatgttttaa tatcttatta aatactgaaa 126660
gcagatcctg atttggaata ggtctcaaga aaggagacta gggtctaatc ctgaattaga 126720
gtctttgctt ataggtaaca aagaatttat gaatttatct cacatttttg cttaagagtt 126780
ctgaatttaa acttccatca aggtcctggg tcccaggtgt tttcaaccta ataattctaa 126840
atattgagtg gttggttgca gtagcttatg cctataatcc cagcgctttg ggaggccaag 126900
gcagaaggat cccttgaccc caggaatgca agaccagcct gggcaacata gtgagacctc 126960
atctctacaa aaaattaaaa agttagctgg acatggtggc ctacacctgt agtcccagct 127020
acttgggaga ttgagatggg aagattgctt gagctgggga ggttgaggct gcagtgagct 127080
gtgatcatgc cactgcatac cagtctgggt ggcagagtaa gactttgtct caaaaaaata 127140
aaaaaataaa aaaaaatgct gagtcaagtc tactgctcct gccagaagag atgactgaag 127200
tgcattacgt agaataataa tggttaaaga aaagctttgc aaaagttccc agaatatata 127260
ctttattggg acaggagaag ctacgtgtgt cgtggtatct ttttactatt ttcttaatct 127320
tataggcctg tgtttctagt caccattaaa ttactacaga tttgtgtttt taatgtaata 127380
tataagtgtt ttggaagggt gagaatattt caaaggtttg aaagttaaaa ctgtgtatga 127440
aagaattgaa aacttggaat ttagatcact tttccattgt gtcatatttc tctgtgacat 127500
ggacatattg aagcatggac atcatgtcca tgcacataaa gcagacaacc cagacaacac 127560
acacatgtgc acaggaacgc tctggaagga tgcaaaccga actgttaagt gttgattcct 127620
gaggaggtga atgggcatgg gttgagagtg agaagaggtt gtaggaagac tttcatatat 127680
tactgtgtac atttctctaa ggtttgaatt ttaaaaaata tattcatgag ttacttttgt 127740
aatgtagaaa tattaatgac tttcctatcc atcagtctgc ctaagcttcc tcttccggtt 127800
caggtagaat gaattggatc agtgttgctc cattttccat tttagcattt tacatttgcc 127860
taaagatatc ttgggatgag ggtatatact ttatcaaatg taagctattt ccaaagtagt 127920
aaatccaaat acttaacaac ttcccagcct aagtaaatga tcagaggctc cgtactcagt 127980
ctcatctaga ctgtggcact gggtgtgaac gtatcaaatg catgtttctc catcaggcag 128040
aagtgagagt aaccatgtgc catggagaag gttgacagac tccctgtgaa gcacttcgaa 128100
gtgacactgg cctctgtgtg cttcagaaga atccagccac ctgctgtgtg gcctgacatt 128160
ttcctttagt ttgtgatggg ccagcagaac tctgttgcca actgttttct gtcctgggtg 128220
cctagccaga ggttctgaaa gtctggagac tttatattgg ctaaacttta ggaacgtcaa 128280
ttacatgtct atctccaaga tgccttcttt tattcaggtg cagctcattg tttcctcttg 128340
agctacactt aagattcttg agcaaaacct aaactgacat ttctccagca atgctctcct 128400
tgagatagaa atgggaaaag taagagcaaa aggaatcttt tgttctcatg tgcatacact 128460
aactcataga aggttaatac ttctatagcc tgtactatta taacaagtat tatatattta 128520
tgatatattt ccttaaagaa aacaaaagca atatagacat ctaaactgtc actggcttat 128580
taagtgtcag tgccagagcc taggagaaaa taaggagcct gtgaattcct tactcgaatc 128640
taaccagagc tgctgtgttt gagagcaagt tttaaaagat tgtatgtaat actaagttta 128700
ttcatctttc acactgagtc ccagcatcac cagatcagta tttgatgcct ggatcaatct 128760
ttattctggg gagtgatgaa gcattgaacc tgctatatgt atagtttgcc gagcgtcggc 128820
atgtgctcct tgtggcccag gcatccctgc atataaggaa taggtacgtt ctcacgagcc 128880
tcacctactt acctccacat ttagccagat tctgggtatt aacatctgct gggaaagagc 128940
atcactacag tagctacaaa taaggtggaa gaagcaaagt atttttctga gaagtactta 129000
aagaatagat gtgtaaattt ctataaacac aagtcttaaa ggaaaatgaa aaaattttac 129060
atttaaataa ctacataaat cattgcctaa ttttaataag aatataactt aatatagctt 129120
gaatggagaa aaggacaact tgcagtcagg gaaagtatta agaaataata tgctcagtct 129180

```
gggcgcggtg gctcatgcct gtaatcccag cactttggga ggccgaggca ggcagatcac  129240
gaggtcaaga gatcaagacc atcctggcta acacagtgaa gccccatctc tactaaaaat  129300
acaaaaaagt agccaggcgt ggtagtgggc gcctatagtc ccagctactc gggaggctga  129360
ggcaggaaaa tgttgtgaac ctgggaggca gatcttgcag tgagtcgaga tcgcgccact  129420
gcactccagc ctgggtgaca gagcaagact ccatctcaaa aaaataaaaa aaataaaaaa  129480
aagaaatatt atgctcaaaa tatatagcaa taagttggaa acttttactt gaataatttt  129540
tacaaaactg accaaagaac aaaaacgtga agaggccaag ttccaagact catgttatgt  129600
aaattgttct aaaacaacat tagcacttaa caaagttgaa aagttaacaa agccaagtac  129660
tgtactaggc ttccaacact aactaagtat aaaattccac agagctggtt ttcttatctt  129720
taaagaaatt tgttggcaag tggtactggt gttaaaaaaa aaaaaaaaaa aggaaatatg  129780
tactgaccaa aatagaaaaa aaatatgaag cacattaaaa gaaaaaaata tatattcctg  129840
aaaaccttgt ataattacag tggcatggtt gggaatgttt ggtctatagt tttaacaatt  129900
aaatccattg aatctggccc cgtaccatcc taaagtttta ttctagattc tctggagttt  129960
gtgattatag atatgtttct aagatttaag taactttcca tgtttatctc cttttatgtt  130020
tgtacataga ataaaaatgt ttctattgtt aagaatatta gagttggacg cagtggctca  130080
cgcctataat cccagcactt tgggaggcca aagcaagtag tttgtttgag cccgggagtt  130140
caagaatggc ctgggcaaca tagtaagacc ccatctctac aaaaaataaa aaattagccg  130200
ggcatagtgg catgtcccag ctacttggga gactaaagtg ggaggatcac tttgagccca  130260
ggaggttgag gctgcagtga gctatgatcg caccactgca ttccagcctg ggcaacaaag  130320
tgagacctgt ttcaaaaata aaaaattggg gtttatctac ttagattttc aataaaaatt  130380
actacttaaa tctttacctg cttgttaatt tcaaaccctt ttctacattt tgatttatct  130440
ttaaatctct ttttgtctca ataaatggga agtatcagga agtcttttta cttgctcaag  130500
gtcatagaga gcttagaacc tggtagtgtc cctctgagcc ccagttcttt ccaacctgcc  130560
aggctgtagg cccaacaatt actcaccact aagaaattat gcttgtgctg tcatggcagt  130620
tgcattggag aaaaggatat ttaactggca aacaaaagtc aggagaatgg ggagattttg  130680
ttcttttgaa atgctagtgt gaagtgctag gcttattttt caaatgccca actcgtattc  130740
ttttcttttc ttttttttct gagagggagt ctcacactgt cgcccaggct ggagtgcagt  130800
ggggcgatct cggctcactg caagctccgc ctgctgggtt gacgccattc tcctgcctca  130860
gcctccgagt aactgggact acaggcgccc accaccacgc ccggctaact ttttttttt  130920
tttttgtatt tttagtagag acggggtttc accgtgttag ccaggatgat cttgatctcc  130980
tgacctcgtg atccgccctc ctcagcctcc caaactgctg ggattacagg cgtgagccac  131040
cgcgcccagc cggccaactc gtattcctaa acgaatcata attttaccat aagaccatag  131100
tttagtgatt gaagaaaaaa tgtaccgaac tgtatgatat gatggtgtca aaaagaacta  131160
acccaatatg aaacagtttt caggagcatg tttcctattt tggtgtcagt ggaccacttg  131220
tgttaaagtt gtgaaaacca acactcctga attccaccca gaactcacac tctgcacctt  131280
cagaggccct cagattgtga gtggctgccc cgaggtgtac taccaacctc cagcttccgt  131340
aggtccgtag gtgtgctagt agggcctagg aaaaacagaa cagatgagga cagtgatgca  131400
tacagctgct ttatctggtc tctccttcct ccctagcctg actgctattg gagggcaccc  131460
tccaggcaca gttctgttca cagcctgctg ctgcccccgt gggccatggt tccaggacgg  131520
ctccatcttc tgtgctttgg gcacattaac ctctcccagc gtcactatct tcatcagcaa  131580
aatggagata acattagtac cacctcataa agttgttatg aggatcacca gtgagataat  131640
caatctaaag tgcttacaac aatgcttggc acttggtaaa cactaaataa atgatagttg  131700
ctattatatg catactttta aaaaacctga tgcttttaaa aatttttttct gctgactagt  131760
gaattgttca gtttttgttg ttgttgttgt tgttgttgtt gtttgagacg gagtctcgct  131820
ctgtcgccca ggctggagtg cagtggcatg atctcggctc actgcaagct ccacccccccg  131880
gattcacgcc attctcctgc cttagcctcc cgagtagctg ggattacagg cgccggccac  131940
cacgcccggc taatttttttt gtatttttag taaagacggg gtttcacctt gttagccagg  132000
atggtctcga tctcccgacc tcatgatctg cccacctcgg cctcccaaag tgctggggtt  132060
acaggcatga gccaccgtgc ccggccatga atggttcagt tttaacaggt tctgtgcctt  132120
aaaaaagtta ttaaattgac tgtttcctcc ttttttgtac ccatcatact ttgaatatat  132180
aactacggca gcatataaac acttcacttg cacttattta tttaaatgtc catctttcca  132240
taacccaagg gtaggaacga aatcttattc attgttgaaa cctctagcac atagcacagt  132300
gcctaccaaa tagtaggcaa aatcaggtgt tcaattctat taactactct aacactgaac  132360
tgaaagtgtt caatggctca aaataatata ataagcctta actctggggt gcttaaattt  132420
atctataatc tctgccagtg aagtatacat agttttaaag gttaaaaaaa aatcaagtgt  132480
ttaatgaatt tgagctgatt gagcactgac cagatacagg atccttaaac ttcataatat  132540
ctagtccaaa gatgaatttt ttttttggta cagattctga cttcaaggga tttgcaggct  132600
gggagggaaa ttaaatatat aaaaagtcat tttcggccag gcgctgtggc tcatgcctgt  132660
aatcccagca ctttgggagg ccgggcaggt ggatcacttg aggtcaggag ttcaagacca  132720
gcctgcccaa catggggaag ccccatctct actaaaaata caaaaattag ctggactttg  132780
```

tggtgcttgc ctgtagtccc agctactcag gaggcttagg cagaagaatt gcttgaacct 132840
gggaagcgga ggttgtaatg agctgagatc acaccactgc actccagcct gggcaacaga 132900
gacagactcc agctcaaaaa caatacgttt ttttaatctt gtcccttaat ggaaatattg 132960
agaaaatgtc taggggaact cgaaggagat gattatagga gttgattatg tatgtaaaat 133020
caaagtgaat gaggcagtgg caggggggaa agggggaacc agtaatgact tagaagtcct 133080
aagcatgttg catggtaatt gtgacatttg cttcctgcga gcggagctga ccttgtggtg 133140
tccgtcctag gtactctcct ttttcttcat tggtctgatg tccatgatga ttcctctgtg 133200
tagcatcttt ggggccctca ttgctgtgtg cctcatcatg ggtctcttcg atggatgctt 133260
catttccatt atggctccca tagcctttga gttagttggt gcccaggatg tctcccaagc 133320
aattggattt ctgctcggat tcatgtctat acccatgact gttggcccac ccattgcagg 133380
taaatataat gattctccag tagttatatt aattcatagt attttctact tcaggtctta 133440
attaagtctc atttatatgt aaaacatatt acaggttatt gattactggt cttttgcttt 133500
tatgtgtgcc ttactggtaa catttttaat aagactagct attaaacagt atgttgaatt 133560
tgctgaagag tccctcttat ccttactggc agctaaaaca ctatataaaa cactttggga 133620
ggctgaggcg ggaagatcac ctgagctcag gagcttgaga ccagcctgga caacatagtg 133680
agacactatc tctacaaaaa ataaaaaata aaaaaattag ccaggcgtgg tggcatgcat 133740
ctgcagtccc agatgcacag gaggctgagg tggggggatt gcttgagccc agaggtcaag 133800
gctgcagtga gctgtgatca taccactgta ccccagccta ggcgaaaccc tatctttaaa 133860
aaaaaaaaaa gttgaagatt ctgaaagatt tctaaatatg tcggcttttg gtaagagatc 133920
atacatcccc ctgaggagta gttgatgaaa aataacatga tagaaaaaac ccaaatgcct 133980
gtaaatagct aactgggtaa gcaaaatgtg gaatatctac agtgcaatat tattcagcca 134040
taaaaaggag taaggtacag aactgatgct accaggtgaa tggacctcaa gaacattatg 134100
ctaagtgaaa gaagccaaac ataaaaggtg acgtggctgg gcatggtggc tcacacctgt 134160
aatcccagca ctttgagagg ccaaggcagg agaatcattt gagctcagga gttccagact 134220
agcttgggta acatagcaag accttgtctc tactaaaaat aaaaaaaatt agcctggcat 134280
ggtgacacat actaatatta gtgtagtccc acctactcca gaggcggagg cgagaggatc 134340
tattgcgcct gggagatcaa ggctgcagag ccgtgatagc accactgcac tccagcctgg 134400
gtgacagagt gagaccctgt ctcaaaataa aaggtcatat atcgtatgat ttatttcatg 134460
tgaaatatcc aaaataggta aattcataca gatagaaaac tggtgactgc caggggctgg 134520
agagaaaggg attgggaagt aactacttaa ttggaatgaa gtgtcctttt ggcatgatga 134580
aaatattttg gaactaaata gaagcagtgg gtgcaaaaca ttgtgtatgt gctaaatgcc 134640
actgaattgt tcactttcaa atggttaatt ctatttcatg tgaattccac ctctagaaat 134700
tctctgaatg atattataaa gatgtcctat gttcagtatt acttgggtgc ctgtgtcaat 134760
tagtattaga agattcagta taaatatgat aaaatttaaa tgtgtgtact gtcatctgtg 134820
tttggcacat gaggtgctca gtgaaggttt agctcaattc attctcatac cctgaactgt 134880
attcctagtc atgaatcaag caagcagagt tactactgag cagaccaaat ccagtttaag 134940
ttcaaaaggt catggtctcc ttataggaga accctccatt ccttcttctc tcccgcaggt 135000
aatatagtat tcactcttct ttctatagcc tacaagctta gttttatagt ttctaagtcc 135060
acagcagatt ttatataagc agcactctag tagtaaatag ttttgctacc agaagtgaca 135120
gttaacagag caagaataag ggaggtaaca ggaaacttct aaagcaggca gaacaccacc 135180
acacagcaac tggcagcaga aagctgctgg ctcaccagga ggctctgagg gcatcatagt 135240
attacagcag aacataatca tgagtgttca taaggaggaa gagttcagga atggtctgat 135300
tccaaagata gaaaacttgt gcatttacag atattgtcta ttttatatat gcagaatgga 135360
tgtacacatt tttatgtgtt ggtacttaat tggacaaatg gtcagctatt tatttaacaa 135420
tgtcagatag atatttagta cagtatattc cctggacatt ttagataact taggttttat 135480
agtataagtt ataagagttt aaaaacacta agataacttt taaaaccatg agtcctggaa 135540
tttgttagag aaattagaaa tgttgagtac cgtaaaagtt ttcagaagca gaacccgaat 135600
atagaatgcc attaataatt cttatatact tacattttct tccaggtact tataaacctt 135660
agtctatttc caaaatgtat ttaacaacct aactttcatc caaattatta aatttaaaat 135720
ttttcttaat tttgaattat gatgttaaat agttttcatg ttatttcata aaaaacaaat 135780
acaccaaaga gtagtttgaa agctgggcat ggtggctcat gcctgtaatc tcagcatttt 135840
gggaggctga ggtgggagga ttgctttagc ctaggagctt gagaccagcc tgggtaatgt 135900
gacaaaaccc catctctaca aaaaatataa aaattagcca ggcatggtag catgcacctg 135960
tagtcccagc tactcaagag gctgagctgg gaggatcaat tgagcttggg aagtcaaggc 136020
ctgcatgagc catgattatg ctaccgcact ccagcctggg tgacacagtg agaccctgtc 136080
tcaaaaaaaa aaaaacgtct acctttctat cttttcaagt aatgtcttgg tatgaaaatc 136140
cagaggactg tactttatga caacgttaaa gatatgagat ctttttcttc ctatcaaaaa 136200
aggtttatat ttcagatctg tttcctaaaa aaaaaaaaaa gtctgatgtc tgagatgtct 136260
gaatcagtcc tgtggctgat ctagaccccc tacagagcca cacttgttct cccttggtga 136320
cagctttttc ccttctcagg gttacttcgt gacaaactgg gctcctatga tgtggcattc 136380

tacctcgctg gagtccctcc ccttattgga ggtgctgtgc tttgttttat cccgtggatc 136440
catagtaaga agcaaagaga gatcagtaaa accactggaa aagaaaagat ggagaaaatg 136500
ttggaaaacc agaactctct gctgtcaagt tcatctggaa tgttcaagaa agaatctgac 136560
tctattattt aatatcttac atacctccac cagactggac ttgctttttg aattttaagc 136620
aagtttcctt tccttttata caaattgcaa atttcatatt tttttaatca catcctagga 136680
atagcacaat aattgggaaa tagaaccctt atcactagaa gaaccatttt ctgccactaa 136740
atatctctga tgtttccatg agtctgaggg cagagactct ggtatatgaa aacatgtctg 136800
aaagtcacat attgtgaaaa tttgaagcta tctcagtaaa aagcagcttt ggaaactgtg 136860
aatgatcttt agcttgtaca aatgtttaaa aatacctcag gctatactga aagggttgca 136920
gtttggttag gagtggaaat attttgtttg ttaatgatgt cttcagttct ggtacctctg 136980
ttttactttc ttatgctctt tggaaacttt ttgcaaaatt taagcctggg ttctagataa 137040
taccagatct acctaaacct caagtctatg ttaaagttgc tttcctgctg ttaaataagc 137100
tatgatatta agatattctg acttgctcca gtgtcaaggg accttctggg agcaggtgct 137160
aacatagtgt tcagaatcaa tatgtgagat gaaaaggatc ccctccagga ggatcctgag 137220
ctgttcagaa atcatttaag tttacagcgt tgttcccttt gcgtttgcag tgcgttttac 137280
tcaagtagcc agaaacaccc cacgtttctg aatttgttta aactgtaaca ataaagtaaa 137340
atagaatgca tgaaagatat tctggcgatt gtaacttaga atttttctga cttctggatt 137400
tgttggcact agaacctgat atttaaacaa agtcttactg agcagctatc aagtggcagt 137460
tacaggcaca aattggtgga ggctggagga tggggagggg agcaaaaccc tttatatttg 137520
tgaagaaaat atctgtagct gatagaaata attgcttaaa ttggtttatg aaattaatga 137580
gtctgaaaag gttaaaagca cttataaaaa gaaccaagtc ctacatttcc agaactttct 137640
ggcaaaaatt tgcactcata ttatttatcc tatgaacatt cccattgttt ttttttgcta 137700
tttatataca gattatcata agaaagctct cagtttgagg acccaaaata aaaccaaagt 137760
catgccatga cccatactca tttacaaaaa caagaacact ttcctctatc cctaaaatta 137820
tgctttagta cttgaggcct ttaaaagtta gtgcttttga ttgtgaagac attcagcaac 137880
ttactttgtc atacatgcag ttgcacctta ccacttctaa tagtgtcata tttcatattc 137940
aggggactta gataatttgc ctgtggatgg ttcttttgca ggaaaaaaaa tctacatttt 138000
gaccatacta ccctttcatg ttcttattat aagcttttag aaaatgattt cattcagtca 138060
tgtccagtta tataaaacgt tactttctca tttttgagaa gttcaacaaa acatactact 138120
aagaccaatc atcaaaccca ctattataaa tgttaatttt ggctgggtaa ggtggcttgc 138180
gcctataatc ccagcacttt gggaggctga ggagggaaga ttgcttgagc ccaggagttt 138240
gagaccagcc tgggcaacat agcaagatcc tgtctctact aaaaataaaa aaaaaaatta 138300
ggccaagcat agtggctcat gcctgtaatc ctagcacttt ggtagtccaa ggcaggggga 138360
tcacttgagc ccagaagttc aagaccagct tgggtaacat aatgagaccc tgtgtctaca 138420
aaaaatttaa aaattagcca ggcatgatgg tgcccacctg tagccccagc tactcagagg 138480
ctgaggtagt ggaaggattg cttgagccta agagatggag gctgcagtga gctatgccac 138540
tgtactctag cctgttcaac tgagcagaac cctgtctgta aaagaaaatc aaaaacaaaa 138600
aataaatgtt aaattttgtt ttaagtttta gcacagactc ccctcaaaac accttctccc 138660
caattttaca gaaagtaatt caaaaatgaa aactttactc tgtaaagacc tctacagtgt 138720
ttttcttttc aaaatttggc tgattttagg aaaaaagtga tcatctgaaa ctaaaagaaa 138780
ttgcttggtt agtttccata ttaaaacagc agtgacaagt atatataact tagatctcag 138840
catatgtgtt tgtatattaa acttcacata tgtagtttc agtttaatgg aatgaatcaa 138900
actggatcta taacactgaa aaagttctat tgtaatagac tcatacggag aatactctgc 138960
tataataata taaaattaag aagaaaaagt ataaacgtaa gatgctaaat tccataaatg 139020
catatttagt actatgtttt ttgtgggaaa agttctaaaa gtttaaaatg cacaaagaaa 139080
atgaaaaata ctaatataaa aatttgtgct ttaatctagt caaactaaat cctttctaat 139140
ttctgaatga agtgttactg ctgcaataaa gtgacctgat aagcctaaat tttttgtgtt 139200
caatccagac acttttctga gagtctgaaa agaatacaga gtcagaactc tgtttttatc 139260
tcctcatcct gtttttgata agactcagaa aattctcaaa ttcgaaaggt tctggcattt 139320
gaggccaaaa aaagcatgaa agggagtaac attccttttt atagatactc tagattggat 139380
actattgtaa cagatggcca agaaacttcc agaaacattt tggttaaatt ttattgcaat 139440
ggatattgct gggatccatc catttaagca gtaatatacc acccagatta ttgatacttt 139500
atgcaagatg tgttcatctc tttgatcata tttacaatgc ttactccata gccctgctac 139560
aagacttaaa att 139573

<210> 4
<211> 1483
<212> PRT

<213> Homo Sapien

<400> 4

```
Gln Met Val Leu Ser Gln Glu Glu Pro Asp Ser Ala Arg Gly Thr Ser
 1               5                  10                  15
Glu Ala Gln Pro Leu Gly Pro Ala Pro Thr Gly Ala Ala Pro Pro Pro
            20                  25                  30
Gly Pro Gly Pro Ser Asp Ser Pro Glu Ala Ala Val Glu Lys Val Glu
        35                  40                  45
Val Glu Leu Ala Gly Pro Ala Thr Ala Glu Pro His Met Val Ser Glu
    50                  55                  60
Glu Pro Ala Arg Thr Glu Ala Gln Pro Gly Pro Ala Pro Ala Pro Pro
65                  70                  75                  80
Pro Gly Pro Ser Asp Val Glu Lys Val Glu Val Glu Leu Ser Met Val
                85                  90                  95
Pro Ser Leu Glu Glu Pro Ala Ala Ala Glu Arg Glu Thr Asn Glu Ala
            100                 105                 110
Gln Pro Pro Gly Pro Ala Pro Ser Asp Asp Ala Pro Leu Pro Val Pro
        115                 120                 125
Gly Pro Ser Asp Val Ser Asp Gly Ser Val Glu Lys Val Glu Val Glu
    130                 135                 140
Leu Thr Arg Ser Thr Gly Asn Gln Gln Glu Pro Pro Glu Pro Pro Glu
145                 150                 155                 160
Gly Gly Trp Gly Trp Leu Val Met Leu Ala Ala Met Trp Cys Asn Gly
                165                 170                 175
Ser Val Phe Gly Ile Gln Asn Ala Cys Gly Val Leu Phe Val Ser Met
            180                 185                 190
Leu Glu Thr Phe Gly Ser Lys Asp Asp Asp Lys Met Val Phe Lys Thr
        195                 200                 205
Ala Ala Trp Val Gly Glu Pro Pro Glu Pro Pro Glu Gly Gly Trp Gly
    210                 215                 220
Trp Leu Val Met Leu Ala Ala Met Trp Cys Asn Gly Ser Val Phe Gly
225                 230                 235                 240
Ile Gln Asn Ala Gly Val Leu Phe Val Ser Met Leu Glu Thr Phe Gly
                245                 250                 255
Lys Asp Asp Asp Met Phe Lys Ala Ala Trp Val Gly Ser Glu Pro Pro
            260                 265                 270
Glu Pro Pro Glu Gly Gly Trp Gly Trp Leu Val Met Leu Ala Ala Met
        275                 280                 285
Trp Cys Asn Gly Ser Val Phe Gly Ile Gln Asn Ala Tyr Gly Val Leu
    290                 295                 300
Phe Val Ser Met Leu Glu Thr Phe Gly Ala Lys Asp Asp Asp Asn Met
305                 310                 315                 320
Ala Phe Lys Ala Ala Trp Val Gly Gln Ser Leu Ser Met Gly Met Ile
                325                 330                 335
Phe Phe Cys Cys Pro Ile Val Ser Val Phe Thr Asp Leu Phe Gly Cys
            340                 345                 350
Arg Lys Thr Ala Val Val Gly Ala Ala Val Gly Phe Val Gly Leu Met
        355                 360                 365
Ser Ser Ser Phe Val Ser Ser Ile Glu Pro Leu Tyr Leu Thr Tyr Gly
    370                 375                 380
Ile Ile Phe Ala Cys Ser Leu Ser Met Gly Met Ile Phe Phe Cys Cys
385                 390                 395                 400
Pro Ile Val Ser Val Phe Thr Asp Phe Gly Cys Arg Thr Ala Val Gly
                405                 410                 415
Ala Ala Val Gly Phe Val Gly Leu Met Ser Ser Ser Phe Val Ser Ser
            420                 425                 430
Ile Glu Pro Leu Tyr Thr Tyr Gly Phe Ala Cys Ser Ser Leu Ser Met
        435                 440                 445
Gly Met Ile Phe Phe Cys Cys Pro Ile Val Ser Val Phe Thr Asp Met
```

```
     450                 455                 460
Phe Gly Cys Arg Arg Thr Ala Val Leu Gly Ala Ala Val Gly Phe Val
465                 470                 475                 480
Gly Leu Met Ser Ser Ser Phe Val Ser Ser Ile Glu Pro Leu Tyr Phe
                485                 490                 495
Thr Tyr Gly Val Val Phe Ala Cys Gln Gly Cys Ser Phe Ala Tyr Gln
            500                 505                 510
Pro Ser Leu Val Ile Leu Gly His Tyr Phe Lys Lys Arg Leu Gly Leu
        515                 520                 525
Val Asn Gly Ile Val Thr Ala Gly Ser Ser Val Phe Thr Ile Leu Leu
    530                 535                 540
Pro Leu Leu Leu Arg Val Leu Ile Asp Ser Val Gly Leu Phe Tyr Thr
545                 550                 555                 560
Leu Arg Val Leu Cys Gly Cys Ser Phe Ala Tyr Gln Pro Ser Leu Val
                565                 570                 575
Ile Leu Gly His Tyr Phe Lys Lys Arg Leu Gly Leu Val Asn Gly Ile
            580                 585                 590
Val Thr Ala Gly Ser Ser Val Phe Thr Ile Leu Leu Pro Leu Leu Leu
        595                 600                 605
Leu Val Gly Leu Tyr Thr Leu Arg Leu Cys Ser Gly Cys Ser Phe Ala
    610                 615                 620
Tyr Gln Pro Ser Leu Val Ile Leu Gly His Tyr Phe Lys Lys Arg Leu
625                 630                 635                 640
Gly Leu Val Asn Gly Ile Val Thr Ala Gly Ser Ser Val Phe Thr Ile
                645                 650                 655
Leu Leu Pro Leu Leu Leu Gly Asn Leu Thr Ser Thr Val Gly Leu Cys
            660                 665                 670
Tyr Thr Leu Arg Ile Leu Cys Gln Ile Phe Met Phe Val Leu Phe Leu
        675                 680                 685
Ala Gly Phe Thr Tyr Arg Pro Leu Ala Thr Ser Thr Lys Asp Lys Glu
    690                 695                 700
Ser Gly Gly Ser Gly Ser Ser Leu Phe Ser Arg Lys Lys Phe Ser Pro
705                 710                 715                 720
Pro Lys Lys Ile Phe Asn Phe Ala Ile Phe Lys Val Thr Ala Tyr Ala
                725                 730                 735
Val Trp Ala Val Ile Phe Met Phe Val Leu Phe Leu Ala Gly Phe Thr
            740                 745                 750
Tyr Arg Pro Leu Ser Lys Lys Glu Ser Ser Ser Ser Phe Ser Arg Lys
        755                 760                 765
Ser Pro Pro Lys Lys Ile Phe Asn Phe Ala Phe Lys Thr Ala Tyr Ala
    770                 775                 780
Val Trp Ala Ser Ile Phe Met Phe Val Leu Phe Leu Ala Gly Phe Thr
785                 790                 795                 800
Tyr Arg Pro Leu Val Pro Ser Ser Lys Glu Lys Glu Ser Glu Asp Ser
                805                 810                 815
Arg Ser Ser Phe Phe Ser Arg Arg Lys Leu Ser Pro Pro Lys Lys Ile
            820                 825                 830
Phe Asn Phe Ala Leu Phe Lys Glu Thr Ala Tyr Ala Val Trp Ala Ala
        835                 840                 845
Gln Gly Ile Pro Leu Ala Leu Phe Gly Tyr Phe Val Pro Tyr Val His
    850                 855                 860
Leu Met Lys His Val Asn Glu Arg Phe Gln Asp Glu Lys Asn Lys Glu
865                 870                 875                 880
Val Val Leu Met Cys Ile Gly Val Thr Ser Gly Val Gly Arg Leu Leu
                885                 890                 895
Phe Gly Arg Ile Ala Asp Tyr Val Pro Gly Val Lys Lys Gly Ile Pro
            900                 905                 910
Leu Ala Leu Phe Gly Tyr Phe Val Pro Tyr Val His Leu Met His Val
        915                 920                 925
Glu Arg Phe Asp Asn Lys Glu Val Met Cys Ile Gly Val Thr Ser Gly
```

930 935 940
Val Gly Arg Leu Leu Phe Gly Arg Ile Ala Asp Tyr Pro Gly Val Lys
945 950 955 960
Lys Ser Gly Ile Pro Leu Ala Leu Phe Gly Tyr Phe Val Pro Tyr Val
965 970 975
His Leu Met Asn His Val Lys Glu Arg Phe Lys Asp Val Asn Asn Lys
980 985 990
Glu Val Leu Phe Met Cys Ile Gly Val Thr Ser Gly Val Gly Arg Leu
995 1000 1005
Leu Phe Gly Arg Ile Ala Asp Tyr Leu Pro Gly Val Lys Lys Gln Val
1010 1015 1020
Tyr Leu Gln Val Leu Ser Phe Phe Phe Ile Gly Leu Met Ser Met Met
1025 1030 1035 1040
Ile Pro Leu Cys Ser Ile Phe Gly Ala Leu Ile Ala Val Cys Leu Ile
1045 1050 1055
Met Gly Leu Phe Asp Gly Cys Phe Ile Ser Ile Met Ala Pro Ile Ala
1060 1065 1070
Phe Glu Leu Val Gly Ala Gln Asp Val Ser Gln Val Tyr Leu Gln Val
1075 1080 1085
Leu Ser Phe Phe Phe Ile Gly Leu Ser Met Met Ile Pro Leu Cys Ser
1090 1095 1100
Phe Gly Ala Leu Ile Ala Cys Leu Ile Met Gly Leu Phe Asp Gly Cys
1105 1110 1115 1120
Phe Ile Ser Ile Met Ala Pro Ile Ala Phe Glu Leu Val Gly Gln Asp
1125 1130 1135
Ser Gln Ser Val Tyr Leu Gln Val Leu Ser Phe Phe Phe Ile Gly Leu
1140 1145 1150
Thr Ser Met Met Ile Pro Leu Cys Ser Val Phe Gly Ala Leu Ile Ala
1155 1160 1165
Leu Cys Leu Ile Met Gly Leu Phe Asp Gly Cys Phe Ile Ser Ile Met
1170 1175 1180
Ala Pro Ile Ala Phe Glu Leu Val Gly Pro Gln Asp Ala Ser Gln Gln
1185 1190 1195 1200
Ala Ile Gly Phe Leu Leu Gly Phe Met Ser Ile Pro Met Thr Val Gly
1205 1210 1215
Pro Pro Ile Ala Gly Leu Leu Arg Asp Lys Leu Gly Ser Tyr Asp Val
1220 1225 1230
Ala Phe Tyr Leu Ala Gly Val Pro Pro Leu Ile Gly Gly Ala Val Leu
1235 1240 1245
Cys Phe Ile Pro Trp Ile His Ser Lys Lys Gln Arg Ala Ile Gly Phe
1250 1255 1260
Leu Leu Gly Phe Met Ser Ile Pro Met Thr Val Gly Pro Pro Ala Gly
1265 1270 1275 1280
Leu Leu Asp Lys Leu Gly Ser Tyr Asp Ala Phe Tyr Leu Ala Gly Pro
1285 1290 1295
Pro Ile Gly Gly Ala Val Leu Cys Ile Pro Trp Ile His Ser Lys Lys
1300 1305 1310
Gln Arg Ser Ala Ile Gly Phe Leu Leu Gly Phe Met Ser Ile Pro Met
1315 1320 1325
Thr Val Gly Pro Pro Val Ala Gly Leu Leu His Asp Lys Leu Gly Ser
1330 1335 1340
Tyr Asp Leu Ala Phe Tyr Leu Ala Gly Ile Pro Pro Phe Ile Gly Gly
1345 1350 1355 1360
Ala Val Leu Cys Leu Ile Pro Trp Ile His Ser Lys Lys Gln Arg Gln
1365 1370 1375
Glu Ile Ser Lys Thr Thr Gly Lys Glu Lys Met Glu Lys Met Leu Glu
1380 1385 1390
Asn Gln Asn Ser Leu Leu Ser Ser Ser Ser Gly Met Phe Lys Lys Glu
1395 1400 1405
Ser Asp Ser Ile Ile Glu Ile Ser Lys Thr Gly Glu Lys Met Glu Lys

1410 1415 1420
Met Leu Asn Gln Ser Leu Leu Ser Ser Ser Ser Gly Phe Lys Lys Glu
1425 1430 1435 1440
Ser Asp Ser Ile Ile Ser Glu Ile Ser Lys Asn Thr Gly Gly Glu Lys
1445 1450 1455
Met Glu Lys Met Leu Ala Asn Gln Ser Ser Leu Leu Ser Ser Ser Ser
1460 1465 1470
Gly Ile Phe Lys Lys Glu Ser Asp Ser Ile Ile
1475 1480